Journal of the
AMERICAN STATISTICAL
ASSOCIATION

UNIVERSITY OF MICHIGAN
SEP 19 1950

SEPTEMBER 1950

Pitfalls in Population Forecasts and Projections ..... Reynold M. Dorn
The Population of the World in the Year 2000 ..... Frank W. Notestein
Population and Resources ..... Joseph S. Davis
Controlled Selection—A Technique in Probability Sampling ..... Leo Goodman and Leslie Kish
A Technique for Analysing Some Factors Affecting the Incidence of Syphilis ..... B. G. Greenberg, John J. Wright, and Cecil G. Sheps
Labor-Force Participation and Employment Opportunities for Women ..... Nedra Bartlett Reiter
Estimating Parameters of Pearson Type III Populations from Truncated Samples ..... A. C. Cohen, Jr.
Response Errors in Census Research ..... Eli S. Marks and W. Parker Mauldin
A Note on the Derivation of Income Estimates by Sources of Income of Persons Making Less than $500 per Annum, 1944-1948 ..... Ulric H. Haber
Estimating Precision of Measuring Instruments ..... H. Fairfield Smith
Horatio Milo Pollock, 1868-1950 ..... Benjamin Meyer
Book Reviews by Palmer O. Johnson, Willard C. Jones, John D. Scott, E. Freund, Louis H. Bean, M. H. Knight, P. J. Lolua, Alvin Mayne, Ro... Bennett, Harold Benjamin, Jean Claire Bowman .....
Briefer Notices .....
Publications Received .....

VOLUME 45 NUMB...

American Statistical Association

Organized November 27, 1839
Incorporated 1841

The American Statistical Association is a scientific and educational organization. Its membership is not confined to professional statisticians but includes economists, business executives, research directors, government officials, university professors, and other persons who are seriously interested in the application of statistical methods to practical problems, in the development of more useful methods, and in the improvement of basic statistical data. Engineers, mathematicians, biologists, actuaries, sociologists, psychologists, and representatives of many other professions are included in the membership of the Association.

Regular membership	$ 8.00*
Student membership	4.00
Introductory membership (for the first dues payment of applicants under 30 years of age)	4.00
Members subscription to Biometrics	3.00
Associate membership in the Biometrics Section	4.00
Contributing membership	25.00

Subscription rate, $8.00 per annum. Prices for back issues available on request.

Additional information about the Association and membership application forms may be secured from the Secretary, 1108 16th Street, N. W., Washington 6, D.C.

The Editors welcome the submission of articles and notes for possible publication in the JOURNAL. Manuscripts should be sent to the Editor, Journal of the American Statistical Association, 1108 16th Street, N. W., Washington 6, D.C. (1) Send two or more copies of your manuscript (preferably in separate envelopes, to avoid danger of loss). (2) Leave one half of the first page of your manuscript blank, to be used for instructions to the printer. (3) Attach to the front of your manuscript a fifty word outline or summary. Authors who wish suggestions about the preparation of manuscripts and charts or information about editorial policies should address their inquiries to the Editor. Book reviews and books for review should be sent to Professor W. Allen Wallis, 207 Haskell Hall, University of Chicago, Chicago 37, Illinois.

---

* $5.00 of which is for one year's subscription to the *Journal of the American Statistical Association*, and 75 cents for one year's subscription to *The American Statistician*.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 45 SEPTEMBER 1950 NUMBER 251

ARTICLES

Pitfalls in Population Forecasts and Projections . . HAROLD F. DORN 311

The Population of the World in the Year 2000 . FRANK W. NOTESTEIN 335

Population and Resources . . . . . JOSEPH S. DAVIS 346

Controlled Selection—A Technique in Probability Sampling . . . . . ROE GOODMAN AND LESLIE KISH 350

A Technique for Analyzing Some Factors Affecting the Incidence of Syphilis . B. G. GREENBERG, JOHN J. WRIGHT, AND CECIL G. SHEPS 373

Labor-Force Participation and Employment Opportunities for Women . . . . . NEDRA BARTLETT BELLOC 400

Estimating Parameters of Pearson Type III Populations from Truncated Samples . . . . . A. C. COHEN, JR. 411

Response Errors in Census Research . . . . . ELI S. MARKS AND W. PARKER MAULDIN 424

A Note on the Derivation of Income Estimates by Source of Income of Persons Making Less than $500 per Annum, 1944–1948 . . . . . ULRIC H. WEIL 439

Estimating Precision of Measuring Instruments . . H. FAIRFIELD SMITH 447

Horatio Milo Pollock, 1868–1950 . . . . . BENJAMIN MALZBERG 452

BOOK REVIEWS

COCHRAN, WILLIAM G., AND COX, GERTRUDE M., *Experimental Designs* . . . . . PALMER O. JOHNSON 454

KAFKA, F., *Statistics without Numbers, a Visual Explanation* . . . . . WILLIAM C. JAMES 456

REICHENBACH, HANS, *The Theory of Probability: An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability. Second Edition* . . . . . JOHN H. SMITH 457

ARLEY, NIELS, AND BUCH, K. RANDER, *Introduction to the Theory of Probability and Statistics* . . . . . JOHN E. FREUND 459

MOSTELLER, FREDERICK, HYMAN, HERBERT, MCCARTHY, PHILIP J., MARKS, ELI S., TRUMAN, DAVID B. (AND COLLABORATORS), *The Pre-Election Polls of 1948* . . . . . LOUIS H. BEAN 461

UNIVERSITIES-NATIONAL BUREAU COMMITTEE ON ECONOMIC RESEARCH, *Problems in the Study of Economic Growth* . . M. M. KNIGHT 464

ROUND TABLE ON INTERNATIONAL STATISTICS, *Problems in the Collection and Comparability of International Statistics* . . . . P. J. LOFTUS 465

ESTEY, JAMES ARTHUR, *Business Cycles: Their Nature, Cause, and Control* . . . . . ALVIN MAYNE 467

INTERNATIONAL ECONOMICS DIVISION, OFFICE OF BUSINESS ECONOMICS, U. S. DEPARTMENT OF COMMERCE, *The Balance of International Payments of the United States, 1946–48* . . . . ROLLIN F. BENNETT 469

BUTTERS, J. KEITH, ASSISTED BY NILAND, POWELL, *Effects of Taxation, Inventory Accounting and Policies* . . . . . HAROLD BENJAMIN 470

MAIR, GEORGE F., EDITOR, *Studies in Population* . . . . . JEAN CLAIRE BOWMAN 471

BRIEFER NOTICES 473

PUBLICATIONS RECEIVED 476

A cumulative Index to Volumes 1–34, 1888–1939, and Annual Indexes thereafter, may be obtained from the Office of the Secretary of the American Statistical Association.

OFFICERS

American Statistical Association

President: S. S. Wilks; *President-Elect:* L. J. Reed; *Vice-Presidents:* H. A. Freeman, D. S. Brady, P. M. Hauser; *Secretary-Treasurer:* Samuel Weiss; *Directors:* G. M. Cox, W. E. Deming, C. H. Goulden, F. F. Stephan, W. L. Thorp, L. L. Thurstone

Members of the Council: C. M. Armstrong, W. S. Brush, W. G. Cochran, D. R. G. Cowan, M. I. Gershenson, M. H. Hansen, H. L. Jones, T. J. Mills, P. R. Rider, David Schneider, J. R. Stockton, E. J. Swan, J. W. Tukey, Sylvia Weyl

EDITORIAL COMMITTEE

W. G. Cochran, *Chairman;* W. Allen Wallis, *Review Editor;* Milton Friedman, Meyer A. Girshick, Jacob Marschak, Douglas Scates, Mortimer Spiegelman

Published Quarterly by the AMERICAN STATISTICAL ASSOCIATION

Publication Office: 450 Ahnaip Street, Menasha, Wisconsin. Editorial Office: 1108 16th Street, N.W.; Washington 6, D.C. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, section 538, P. L. & R., authorized March 25, 1936. Entered as second class matter at the post office at Menasha, Wisconsin.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 251 SEPTEMBER 1950 *Volume 45*

PITFALLS IN POPULATION FORECASTS AND PROJECTIONS*

HAROLD F. DORN

This paper reviews the development of demography, and especially of population forecasting, and discusses the current demographic situation. The merits and limitations of several methods of making population predictions, and errors in demographic thinking are examined. The complexity of the problem of making forecasts either for short periods or long-time trends with the techniques now available is pointed out.

EXPERTS ON population have projected their curves into the future and the outlook is startling. Manufacturers who try to estimate future markets have been expecting a population of 140 million by 1940, but the calculations of our contributors . . . show . . . that hardly more than 132 or 133 million are to be expected by 1940." . . .

"As our statisticians look further into the future, they see possibilities of still greater declines in growth with the probability of a stationary population. They show that we shall probably attain a population between 145 and 190 million during the present century with the probability that the actual population will be nearer the lower figure than the higher. Such a prospect is radically different from that predicted a generation or even a decade ago" [1].

These words written by President Hoover's Research Committee on Social Trends in 1933 marked the beginning of a new era in the development of demography. Population forecasts were not new in the United States. Since the earliest days of the nation, geologists, business men, writers, economists, statesmen and even presidents had published predictions concerning the growth of the population of the United States. But no single group of these forecasters ever had been generally recognized as having special competence to prognosticate demographic

* Revision of a paper presented before a joint meeting of the American Statistical Association and the Population Association of America in New York City on December 30, 1949.

changes. As a result, persons from all walks of life felt free to peer into the crystal ball of the future and to chart the growth of the nation's population.

The words of President Hoover's Research Committee on Social Trends recognized a particular group, called demographers, and gave endorsement to their special ability to forecast population changes. This endorsement was widely accepted by the public, by business men, by governmental agencies and by fellow scholars and scientists.

It was an auspicious beginning. Seven years later the Director of the Bureau of the Census announced that the population of the United States on April 1, 1940 numbered 131,669,275 persons. Allowing for underenumeration, officially recognized by the Bureau of the Census, demographers nearly a decade prior to this date had predicted the total population more accurately than the Bureau of the Census had been able to count the actual population! In the words of Irene Taeuber in an article reprinted with approval by the editors of *Population Index*, "With improved data, new techniques, and the precise measurement of the demographic transition that was occurring, demography tended to become science rather than literature" [2].

Confidence in the reliability of estimates of future population became widespread. O. E. Baker undoubtedly expressed the opinion held by many demographers and non-demographers alike when he said, "The population of the United States ten, twenty, even fifty years hence, can be predicted with a greater degree of assurance than any other economic or social fact, provided the immigration laws are unchanged" [3].

This belief gained wide acceptance in spite of the fact that estimates frequently were prefaced with disavowals of predictive accuracy. The following statement by Thompson and Whelpton is typical:

"Some care must be exercised in the use and interpretation of these estimates. It is to be emphasized that they are not predictions of future population size, nor are they to be assumed to indicate the probable sex and age structure. They are, strictly speaking, merely statements of what the size and the sex, age, color and nativity composition of the population would be at specified future times if birth rates, death rates, and immigration were to follow certain specified trends.

While it is true that the fertility and mortality assumptions have been chosen with regard to the indicated trends of vital rates, the purpose of the estimates is to show the approximate range within which future populations would fall under the influence of current trends, rather than to arrive at a single most probable figure" [4].

But most readers obviously assumed that these sentences were merely a polite bow to scientific caution and did not take them very seriously. Nor can they hardly be blamed. For why should demographers be continually revising population forecasts with ever increasingly elaborate assumptions if they themselves did not believe that the forecasts had some validity? And were not demographers writing papers and making speeches which either implicitly or explicitly accepted the validity of the forecasts? It seemed incredible that demographers were merely doing sums in arithmetic for self entertainment; they must be serious and really believe that the projected population estimates were in fact reasonable forecasts or predictions of the future.

So overwhelming was this belief that economists, governmental officials and business men accepted the projections as being substantially correct. This acceptance was buttressed by the developments of population change during the decade of the 1930's.

On January 31, 1941 the Bureau of the Census reported, "If the 1930 birth and death rates had continued, the population of the country would more than reproduce its numbers by about 11 per cent per generation. . . . In 1940, however, as a result of changes in fertility and mortality during the decade, the population was no longer maintaining its numbers. As a matter of fact, if the 1940 birth and death rates continue, the population would in the long run fail to reproduce itself by about 4 per cent per generation" [5].

But, as a matter of historical fact, the 1940 birth and death rates did not continue unchanged. The men and women who bear children refused to have their personal relationships regulated by governmental press releases. By 1942 the birth rate had increased 20 per cent since 1939. Concerning this increase, the Metropolitan Life Insurance Company stated, " . . . the conditions of 1942, continued unchanged into the future, would lead us to a population eventually increasing at the rate of 18 per cent per generation" [6]. Did this really mean that the hypothetical John Q. Public, who in January 1941 was doomed to become an ancestor without progeny, now could look forward to a long, if not numerous, line of descendants?

In my analysis of this change in the birth rate in September 1942 I pointed out that the increase had resulted principally from a larger number of first and second births and concluded, "It is not possible to permanently maintain a population by an increase in the number of first and second births alone, especially when the number of births of the fourth order and higher is decreasing. For this reason the computation of net reproduction rates based on the number of births registered

during one year not only is unwise but may also be misleading. . . . "

"The apparent increase in the birth rate since 1933 is the result of an increase in the number of first and second births and, consequently, does not represent an increase in the lifetime fertility of the women of childbearing ages. It is doubtful that this increase will continue beyond 1942. After that date a decline seems inevitable, at least for the duration of the war. If such is the case, the potential decrease of the population, somewhat prematurely announced by the Bureau of the Census in 1941, will become a reality" [7].

After increasing slightly during 1943, the birth rate started declining but by the end of 1945 still was about 13 per cent higher than in 1939. Beginning in 1946 the trend was abruptly reversed; the 1947 birth rate exceeded the 1939 rate by nearly 50 per cent. This was the highest birth rate since possibly 1920. Concerning this, Whelpton said in a report of the National Office of Vital Statistics dated October 7, 1948, "It is just as improbable that the high rate of 1947 will remain in effect as it was that the low rates of 1933 would do so. Both were the results of unusual conditions—the great depression of the early 1930's and the demobilization and high prosperity following World War II" [8].

During 1948 the birth rate declined about 4 per cent. By now some demographers were becoming uneasy. The April 1949 issue of *Population Index* stated, "Until recently the course of population development in Western nations was generally believed to be well charted and understood. This is now a matter of some doubt" [9].

The disquietude of consumers of population forecasts was vigorously expressed by Joseph S. Davis of the Food Research Institute at Stanford University before the American and Western Farm Economics Association in Laramie, Wyoming on August 20, 1949. With all the righteous indignation of a trusting wife who has just discovered the activity of a philandering husband, Professor Davis charged, "I am ashamed that, like most of my fellow social scientists, I have so long accepted the conclusions of the population specialists with naïve faith. . . it is disheartening to have to assert that the best population forecasts deserve little credence even for 5 years ahead, and none at all for 20–50 years ahead. . . . Population forecasting is *not* a simple matter. Available techniques do *not* permit reliable prediction to be made for 5, 10, 20 or 50 years ahead. The best may be *far* wrong. Our net reproduction rate is *not* near unity, but has been well above it ever since 1940. It is *not* reliable as a basis for prediction. There is *no* assurance of any peak population at any future date. The age structure of the population does

not 'inherently' point to cessation of growth and eventual population decline. Our major population problems are not prevention of such decline. There is no adequate basis for expecting the fertility rate, or the crude birth rate, to drop to or below the level of the early 1930's and to remain at that low level . . . planning for food, agriculture, industry, schools, et cetera, can *not* be safely done on the basis of supposedly expert population forecasts . . . . If we continue to build on the crumbling foundations I have described, we shall have no excuse for consequent errors in our own work" [10].

With this devastating indictment Professor Davis arraigned demographers before the bar of judgment not only of fellow social scientists but of the entire nation. Are they guilty or not guilty? Before rendering final judgment let us first look at the development of demography, and more especially population forecasting, in historical perspective and then examine briefly the current demographic situation.

EARLY FORECASTS

As I pointed out previously, forecasting the future population of the United States is not a recent development. Benjamin Franklin, as early as 1751, estimated that the population of the colonies was doubling about every 20 years due to natural increase alone [11]. The early forecasters had almost no basic demographic data beyond the total size of the population available as a guide and even these meager data covered only a very few years. It is not surprising then that forecasts prepared during the early years of the nineteenth century assumed an indefinite increase in population at a geometric rate. Several presidents, including Thomas Jefferson and Abraham Lincoln, in the days when presidents undoubtedly prepared their own papers, advised Congress that the population of the nation would continue to increase at a rapid rate. In his second message to Congress, Lincoln thought that the total population very likely might be as large as 102 million in 1900 and 217 million around 1925 [12].

By the middle of the nineteenth century the belief that population could continue to increase according to the Malthusian Law, $P = Ce^{rt}$, was being questioned. The Compendium of the United States Census for 1850 summarized the projections of various forecasters and presented a number of new estimates based on a decrease in the rate of growth.

Several of these early predictions proved to be surprisingly accurate for relatively long periods of time. The estimates of Bonynge published in 1852 did not deviate from subsequent census enumerations by as

much as 3 per cent until 60 years later in 1910 when his estimate exceeded the census enumeration by 5 per cent [13]. His estimates for later dates, however, have considerably exceeded census enumerations and the estimate of 703 million for 2000, seems likely to be at least 3 to 4 times the possible population on that date.

Another remarkably accurate series of predictions were those prepared by Gannett for the National Conservation Commission in 1909 [14]. These deviated by not more than 3 per cent from the census enumerations from 1910 through 1940; his estimate for 1950, 150 million, will miss the census count probably by less than one per cent. The estimate of 249 million for 2000 undoubtedly is considered by most demographers as at least 25 per cent too high. But as will be pointed out later it is viewed as being exceedingly probable or even somewhat too low by some economists. It should be remembered though that the year, 2000, is nearly a century after the preparation of the original estimate.

But as Gannett so aptly remarked, "Several predictions of our population have been made at various times by different and more or less complicated and elaborate methods. One or two of these predictions have by accident hit very near the truth, as made known by later censuses, but all have been finally wide of the mark" [15]. Moreover these estimates were all wide of the mark because they assumed a much larger increase in population than actually took place. Evidence is abundant to demonstrate that the slower than expected increase of population did not result from population increase outrunning means of subsistence. Something had happened to the Malthusian Law.

By the end of the nineteenth century, some persons trained in the more rigorous disciplines of mathematics and the natural sciences began to question the value of qualitative wisdom or the mechanical extrapolation of trend curves such as the Malthusian Law as methods of prediction. One of the first of these was Pritchett writing in the *Quarterly Publications of the American Statistical Association* in 1891 and again in the *Popular Science Monthly* in 1900 [16, 17]. He observed that the Malthusian Law could not accurately describe the growth of population in the United States and that, in general, the Law was irrational since with indefinite increase in time population also would increase without limit, which is manifestly impossible.

Pritchett restated the general law of population growth as follows: When not disturbed by war, famine, immigration, etc., population increase goes on at a constantly decreasing rate. It should be possible to express this law by means of some rational mathematical function.

"And it does not in the least diminish the value of such a mathematical formula, for the purpose of prediction, that it is based upon no knowledge of the real causes of the phenomena which it connects together" [17].

Examining the growth of population from 1790 to 1880 Pritchett observed that it could be reproduced very closely by a third degree power series, $P=A+Bt+Ct^2+Dt^3$. The fact that this function had no upper asymptote and that if extrapolated backward in time P would become zero or even negative apparently did not seriously disturb Pritchett. So firm was his faith in the ability of rational mathematical functions to describe demographic phenomena that he prepared estimates for ten centuries in advance.

His faith brought little credit to the American Statistical Association which published his forecasts nor did it positively advance knowledge of how to forecast population growth. The estimate for 1910 was only 3 per cent too high but the error increased rapidly thereafter. The estimate for 1950, 191 million, will be about 27 per cent in excess of the census enumeration. The estimates of 386 million for 2000 and of 41 billion for 2900, do not appear realistic at this time.

LOGISTIC CURVE

Attempts to predict population growth by means of some rational mathematical function involving population and time reached their zenith in 1920 in the publication by Pearl and Reed of a paper in which they derived an empirical curve which conformed to certain reasonable postulates concerning the growth of populations [18]. This was an independent discovery of a function which much earlier in 1838 had been suggested by the Belgian mathematician, Verhulst, but which had remained unnoticed undoubtedly due to the absence of sufficient reliable demographic data by which its validity could be tested [19]. This curve, $P=K(1+e^{a+bt})^{-1}$ now known as the logistic was placed on a firm rational basis by Lotka shortly after the publication of the paper by Pearl and Reed [20].

I do not believe it essential for the purposes of this discussion to develop the rational basis of the logistic. It will be sufficient to state that it is in harmony with our rational ideas about population growth. Moreover, there can be no doubt concerning the theoretical validity of the logistic as a description of many diverse types of growth phenomena. But the question confronting us is, does the logistic represent the past growth of the population of the United States with such reliability that it can, with confidence, be used to predict the future?

TABLE I

ENUMERATED AND ESTIMATED POPULATION OF THE UNITED STATES
(in millions)

Year	Census[1]	Pearl[2]-Reed[3]		Dublin[4] 1931		Scripps[4] 1928	Scripps[5] 1931	Scripps 1933[6]		Scripps 1935[7]			Scripps 1943[8]			Scripps 1947[9]			Census[10] 1949
		I	II	High	Low			High	Low	High	Medium	Low	High	Medium	Low	High	Medium	Low	
1900	76.0					(Jan. 1)	(Jan. 1)				(April 1)						(July 1)		(July)
1910	92.0		91.4																
1920	105.7	107.4	106.1																
1930	122.8	122.4	120.1			123.6													
1940	131.7	136.3	132.8	131.0	131.0	138.3	132.5	134.5	132.5	132.6	132.0	131.2							
1941	133.2																		
1942	134.7																		
1943	136.5																		
1944	138.1																		
1945	139.6									139.2	137.0	134.1	138.7	138.5	138.3				
1946	141.2																140.8		
1947	144.0																142.2		144.0
1948	146.6																143.3		146.2
1949	149.2																144.5		148.2
1950		148.7	143.8	139.0	139.0	151.6	139.8	148.5	140.5	146.1	141.6	136.2	145.0	144.4	143.0	148.0	146.0	144.9	149.9
1960		159.2	153.0	147.0	146.0	162.7	143.9			159.5	149.4	137.1	156.5	153.4	147.7	162.0	155.1	149.8	160.0
1970		167.9	160.4	151.0	148.0	171.5	144.6		146.0	172.8	155.0	134.0	167.9	160.5	148.7	177.1	162.0	151.6	
1980		174.9	166.3	154.0	148.0		142.9	190.0	145.0	185.8	158.3	127.6	179.4	165.4	145.8	in-	Increase	decline	
1990		180.4	170.8	154.0	145.0		declines		declines		declines	declines	189.4	167.1	138.9	crease	until	there-	
2000		184.7	174.3	152.0	140.0	186.0	there-		there-		there-	there-	198.7	166.6	129.1	there-	about	after	
		ultimate 197.3	ultimate 184.0				after		after		after	after		decline	decline	after	2000 then decline		

[1] Enumerated population, 1900–1940; calculated as of July 1, 1941–1949.
[2] *Science*, 92, 486–488, 1940.
[3] *Problems of Population*, 1932, pp. 115–125.
[4] *The Am. J. Sociol.*, 34, 253–270, 1928.
[5] *Problems of Population*, 1932, pp. 77–86.
[6] *Recent Social Trends*, p. 48.
[7] *J. Am. Stat. Ass.*, 31, 457–473, 1936. (High is based on high fertility, low mortality, 200,000 net immigrants annually beginning 1940; medium is based on medium fertility and mortality with 100,000 net immigrants annually beginning 1940; low is based on low fertility, high mortality, no net immigration.)
[8] National Resources Planning Board, 1943, p. 29. (High is based on high fertility, low mortality, no net immigration; medium is based on medium fertility and mortality, 100,000 net immigrants annually beginning 1945; low is based on low fertility, high mortality, no net immigration.)
[9] Bureau of the Census, 1947. (High is based on high fertility, low mortality, 200,000 net immigrants per year beginning 1945; medium is based on medium fertility and mortality, 100,000 net immigrants annually beginning 1950; low is based on low fertility, high mortality, no net immigration.)
[10] Bureau of the Census, Series P-25, No. 18, 1949.

Pearl and Reed fitted their first logistic to the population of the United States from 1790 to 1910 before the results of the 1920 census were available. They forecast the 1920 and 1930 populations with an error of less than one per cent. The forecast overestimated the 1940 population by 3.5 per cent and it will underestimate the 1950 population by about one per cent. (These forecasts, together with others discussed below, are presented in Table I.) The projections are accurate enough to satisfy all except the most captious critics and set an example of forecasting ability which might well be the envy of other social scientists.

But was this accuracy really due to the discovery of the natural law of population growth and the superior ability of demographers to foresee the future or was it merely a fortuitous historical event such as the earlier forecasts prepared during the nineteenth century which were mentioned above?

As soon as a preliminary count of the 1940 census enumeration was announced Pearl and Reed published a comparison of this count with the estimate by Logistic I [21]. Even though the estimate exceeded the census count only by about 3.5 per cent Pearl and Reed did not have the courage of their convictions. Apparently they agreed in part at least with Thompson and Whelpton who earlier had written, "An *S* shaped curve has been fitted to the past population of the United States by Pearl and Reed, and prolonged to indicate the future population. To the authors, it seems probable that these estimates will prove too high due to restricted immigration and lowered birth rates" [22].

Pearl and Reed were prepared for this contingency. Their earlier investigations of the ability of the logistic to describe the growth of a great variety of populations had led to the enunciation of the principle, that any population can be expected to continue to follow, in its later growth, the same logistic curve it has followed in its earlier growth, only if no serious changes have occurred in conditions governing previous growth. If changes do occur, a new logistic should be computed.

In spite of the fact that the 1940 forecast met rigid mathematical tests for a good forecast in that it deviated from the census enumeration by less than twice the standard error of forecast, Pearl and Reed decided to compute Logistic II from census data for the entire period 1790–1940. The fit to past growth was remarkably close and resulted in an estimated 1940 population less than one per cent greater than the census count. Pearl and Reed did not express a clear choice for either Logistic I or Logistic II but suggested that the relative merits should be decided by future events. However, they seemed to have a slight

preference for the fledgling Logistic II over the time tested Logistic I which had remained tried and true for 30 years.

Alas, their faith was ill rewarded. In 1950, Logistic I will underestimate the census count by about one per cent but the newer Logistic II will fail in its first test since it will underestimate the census count by about 4 per cent. This may not appear large when compared with predictive errors of many other social, economic and biological phenomena but it fails to meet the mathematical requirements for a good forecast since it will deviate by about twice the standard error of forecast from the census count and what is worse this error will become progressively greater in the future. No, here was not a mathematical law which could mysteriously, within its parameters, contain the forces governing past population growth in such a manner that, following the demographer's touch, the future would be accurately portrayed. Just as for the Malthusian Law, something had happened to the logistic law.

USE OF NET REPRODUCTIVE RATES

A method of analysis of population growth, different in conception from those discussed so far in this paper, had gained widespread acceptance by the early 1930's. In 1911, Sharpe and Lotka had shown that a population continuously subject to a fixed set of fertility rates for women of each age and a fixed set of mortality rates for each age, in the absence of migration, ultimately would assume a stable age distribution [23]. The ultimate birth rate, death rate and rate of natural increase, therefore, could be computed and compared with those prevailing at any time prior to stabilization. These ultimate rates were called true birth and death rates and their difference, the true rate of natural increase.

These concepts were extended by Lotka in subsequent papers, but were first brought forceably to the attention of demographers by the publication of a paper by Dublin and Lotka in the *Journal of the American Statistical Association* in 1925 in which these concepts were applied to the analysis of population growth in the United States [24]. They pointed out that as a result of high fertility and mortality rates in the past, the current age distribution of the population was unusually favorable to a high crude rate of natural increase. However, if the existing fertility and mortality rates should continue unaltered, in the absence of migration the age distribution of the population would change in such a manner that the annual rate of natural increase would be reduced about 50 per cent, from 11.1 to 5.5 per 1,000.

Lotka also showed how to compute the ratio of total births in two generations. This had been first shown by Böckh in 1884, but had been subsequently overlooked [25]. This ratio, now known as the net reproduction rate, was given wide publicity by the publications of Kuczynski beginning in 1928 [26].

Net reproduction rates for all possible segments of the population were computed and published. When the requisite birth and death statistics were lacking, demographers invented ingenious substitutes for the net reproduction rate. Never before had the procreative effort of a population been so closely studied.

As demographers kept a sensitive finger on the fluttering demographic pulse of the population, their regular bulletins pointed more and more to the possibility that the population of the United States was but a withering branch of a decaying ancestral tree. Not only was the true or potential rate of population growth rapidly declining to a level which would eventually result in a decrease in the size of the total population but also the actual excess of births over deaths was becoming smaller and smaller. In 1939 I pointed out that during 1935 and 1936, the first two years for which vital statistics then were available on a residence basis, 5 of the 93 cities of 100,000 or more population and 14 cities of 50,000–99,999 population actually had more registered deaths than births [27].

It is not surprising that population forecasts based on the true rate of natural increase reflected the pessimism of the time. In 1931, Dublin presented two projections of the population of the United States before the General Assembly of the International Union for the Scientific Investigation of Population Problems in London [28].

The most optimistic forecast assumed that the birth rate would continue to fall until it reached a level which would support a stationary population under the best mortality conditions foreseeable. An expectation of life at birth of 70 years by 1970 seemed the maximum possible. Accordingly, Dublin assumed that the true birth rate and death rate would be equal and stabilized at 14.3 by 1970 and that they would continue at that level until 2100.

With these assumptions the population of the United States faced a dismal future. In the absence of migration, it would reach a maximum of 154 million between 1980 and 1990 and then decline to 140 million by 2100. Even so Dublin reported, "the supposition on which this first estimate is based seems altogether too optimistic."

His second "more reasonable" estimate assumed that the true birth rate would fall to a level of 13 per 1,000 by 1970 and then gradually

decline until it reached a level of 10 per 1,000 by 2100. With the same mortality assumption as before and in the absence of migration, a maximum population of 148 million would be reached by 1970 followed by a decline to 140 million by 2000 and to 76 million by 2100.

These predictions are remarkable not so much for their specific numerical values as for the fact that a prominent American demographer had stated publically that the population of the United States almost certainly would decline in the very near future. The most pessimistic of previous forecasters had assumed merely that the rate of increase would approach zero at some distant time. Moreover, Dublin expressed the opinion that "my extreme prediction for 2100 may turn out to be conservative after all."

It is hardly necessary to ask, how have these forecasts stood the test of time? The census enumeration in April 1950 will report a population larger than the maximum resulting from Dublin's second forecast and within 3 or 4 million of the maximum forecast by his first optimistic projection. No, the true rate of natural increase and the net reproduction rate, in spite of their theoretical importance, did not hold the key to the future. In fact they seemed to obscure the future. Just as with the Malthusian Law and the logistic curve, something was wrong with the net reproduction rate.

ESTIMATES OF SCRIPPS' FOUNDATION

The most authoritative and widely accepted series of population estimates have been those flowing from the tireless industry of Whelpton of the Scripps Foundation. His procedure which he terms the analytical method, "consists of (1) analyzing mortality, fertility, and migration trends in different population segments, (2) framing hypotheses regarding the future trends of these factors by observing their previous behaviour and appraising the influence of changing industrial, social, and legal conditions, and (3) building up successive hypothetical populations, beginning with the last actual population, by applying hypothetical factors to different population segments. The results thus show the composition as well as the size of the indicated future population" [29].

The first of these predictions was published in 1928 [30]. I shall call the numerical results of Whelpton's computations, predictions or forecasts for convenience in discussion. However, he has steadfastly maintained that he was not predicting the size of the population at some future date. "No claim is made that the Scripps Foundation estimates represent a law of population growth. They are simply the results of an

empirical process. . . . These estimates represent simply what will happen under certain conditions of immigration, birth-rates, and death-rates, conditions that are believed to be reasonable, based on the experience of recent years" [31].

I shall not digress to discuss whether a rose is a rose. Regardless of Whelpton's protestations, the results of his computations have been widely regarded and used as forecasts of future population change not only by fellow demographers but by other social scientists and by governmental officials and businessmen.

The 1928 projection missed the 1930 enumerated population by less than one per cent but overestimated the 1940 census enumeration by about 7 million or somewhat more than 5 per cent. In 1950 though, it will again be within one per cent of the census population. The population would continue to increase at a decreasing rate, reaching 186 million by 2000. Although prepared by an entirely different method, these estimates were closely similar both in trend and absolute size to those resulting from Logistic I prepared by Pearl and Reed.

Concerning these estimates Whelpton said, "It is true that striking medical discoveries may cause the population to vary upward from these predictions. More likely, however, wars or a greater practice of of birth control may cause a variation downward" [32].

Additional study apparently convinced Whelpton that the last sentence was true, for three years later before the same sessions of the International Population Union at which Dublin presented the estimates mentioned above, he read a paper which made the pessimistic Dublin appear relatively optimistic [33]. The population of the United States now was scheduled to reach a maximum of 144.6 million by 1970 and to decline rapidly thereafter.

Whelpton explained this radical change in point of view as follows: "The fact that the present estimates indicate a population for 1975, which is about 31,000,000 below that indicated by the estimates prepared four years ago, serves to illustrate the cumulative effect during a 45 year period of the acceleration in the decline in the birth-rate from 1927 to 1931, since the same method of estimating was followed in both cases and little change was made in the trends of death-rates and immigration. Perhaps the actual course of population growth will lie between the two extremes, although at present it appears as though the lower estimates are the more probable, and that the maximum during the century will not exceed 150,000,000, even though surpassing the 142,600,000 mark resulting from our present calculations" [34].

The population projections of the Scripps Foundation were given an

aura of authority with their acceptance by President Hoover's Research Committee on Social Trends. Here was the stamp of approval of fellow social scientists who presumably would be expected to be most critical. Moreover, the Committee had at least quasi-official standing.

The report of the Committee published in 1933 contained two alternative forecasts of population prepared by the Scripps Foundation [35]. The low estimate did not differ materially from the forecast presented by Whelpton before the International Population Union in 1931. The high estimate was somewhat less than Whelpton's 1928 forecast up until 1950 but after that date it assumed a less rapid decline in the rate of growth so that the population would number 190 million by 1980.

The authors placed very little confidence in the high estimate. Concerning it, Thompson and Whelpton said, "It is believed by the authors, however, that the actual population will be considerably nearer the minimum than the maximum figure, especially by 1980" [36]. They preferred a forecast obtained as a weighted average of the two estimates, but giving increasingly greater weight to the lower up to 75 per cent for 1980. On this basis the future population would be about 143 million in 1950, increasing slowly thereafter to about 157 million by 1980. The future still looked dark.

The interest and concern of various governmental agencies about the effects of the gloomy population outlook upon the nation found expression in a report, *The Problems of a Changing Population*, published by the National Resources Committee in 1938 [37]. Revised estimates of future population changes, prepared by the Scripps Foundation, were included in this report and strongly influenced the Committee's recommendations.

The estimates prepared for this report represented a full-blown shot gun approach to population forecasting. All possible combinations of the various assumptions concerning the trend of fertility, mortality and migration gave 18 series of possible future population change. This wide variety of projections emphasized the position Whelpton always had maintained but which few people really believed, namely that his projections were not forecasts of actual populations but were intended to show the inevitable consequences of the continuation of certain assumed trends in fertility, mortality and migration. Only three series will be mentioned, here, those selected by Whelpton to represent high, medium and low estimates.

The low series was more pessimistic than any previously published. The population would reach a maximum of about 137 million between

1955 and 1960 and thereafter decline rapidly [38]. The high series resulted in forecasts slightly lower than the high estimates prepared for the President's Research Committee on Social Trends. The total population would continue to increase, though at a decreasing rate, and by 1980 would number about 186 million. Whelpton did not consider either of these very probable for he said, "At the present time it seems to us that the medium series will be closer to what happens than the high or low, though probably somewhat too high. According to this series, the population will reach its maximum of about 160,000,000 soon after 1980, and then begin to dwindle numerically" [39].

This opinion was shared even more strongly by the Committee on Population Problems for it did not even include the high estimate in the body of its report. The highest estimate shown in detail by the Committee was Whelpton's medium series which assumed medium fertility, medium mortality and net immigration of 100,000 annually after 1940. As pointed out above these assumptions would result in a declining population before the end of the century. The possibility that the population of the United States, before the end of the century, would reach a maximum from which it would decline was becoming almost a truism. Even a governmental agency could endorse this belief without criticism. A slight modification of the medium estimate (assuming no net immigration) was republished, with apparent approval, by the Bureau of the Census in 1941 [40].

The onset of the war, a somewhat more rapid decline in mortality than had been anticipated and the continuation of the upward turn in the birth rate which had started in the middle 1930's led the National Resources Planning Board to request the Scripps Foundation to revise the 1937 forecasts. The revised estimates published in 1943 were higher than the 1937 figures but under most combinations of assumptions concerning the trend of fertility and mortality, a gradual slowing down and an eventual cessation of population growth was projected [41]. As Thompson and Whelpton said, "the outlook for the population of the United States remained much the same."

In spite of war casualties, the population continued to increase so rapidly that by 1945, it exceeded the highest estimate made only 3 years previously. Although the Scripps Foundation saw no reason for materially altering its long range projections of fertility and mortality trends, it was apparent that the "baby boom" had made obsolete the base population on which future projections rested. Accordingly, a set of revised estimates were prepared and published in 1947 under the sponsorship of the Bureau of the Census [42].

Whelpton summarized these by saying, "The outlook after 1950 is for a continuation of the long time decline in population growth, both in absolute number and rate. Moreover, there is a strong possibility that within a few decades the population will reach its maximum size and will begin to decrease unless heavy immigration is resumed" [43]. This view also must have been shared by the Bureau of the Census for it had selected the forecasts based on medium fertility and mortality and no net immigration for publication in advance of the complete report [44]. According to this set of assumptions, a maximum population of some 165 million would be reached about 1990 after which a decline would occur.

Because of the interest in such figures, the report contained a series of annual forecasts for 1946–1949. By the time the report was published in the latter half of 1947, the actual population at that time, as subsequently calculated by the Bureau of the Census, exceeded the number forecast in the report by nearly two million.

With population increase outstripping the speed of governmental printing presses, heroic action clearly was called for. On February 14, 1949 the Bureau of the Census issued revised forecasts of population for July 1, 1948 to 1955 [45]. The forecast for July 1, 1950, was nearly 2 million greater than the highest and five million greater than the lowest 1947 forecast for that date. As of December, 1949, it is likely that this revised "revised" estimate may be nearly two million too low.

REASONS FOR PAST FAILURES

Demography, is it science or literature? It is no answer to the record I have just sketched to say as did the editors of *Population Index* in their review of the 1947 forecasts prepared by the Scripps Foundation and the Bureau of the Census, "the revised estimates are not predictions. They are projections . . . " [46]. Predictions, estimates, projections, forecasts; the fine academic distinction among these terms is lost upon the user of demographic statistics. So long as numbers which purport to be possible future populations are published they will be regarded as forecasts or predictions irrespective of what they are called by the demographers who prepare them.

What is wrong with the methods of forecasting population I have just described? Must they be discarded as unsound and should demographers as Professor Davis suggests, return to qualitative wisdom as a forecasting method? [47]

It is hardly necessary to mention why the Malthusian Law failed. It is inherently irrational since it assumes that population will increase

indefinitely with indefinite increase in time. But what of the logistic curve, popularly known as the law of population growth?

So far as I have been able to judge neither Pearl nor Reed, at least as soon as the first flushes of parental pride had subsided, maintained that the logistic was anything more than an empirical curve which described exceptionally well the past growth of the population of the United States. Its standing as a law of population growth was critically examined by Wilson and Puffer in 1933 in a publication apparently little known to demographers with the conclusion, "If . . . the statement that the logistic . . . affords a rational law to such an extent as to permit the extrapolation of the curve for forecasting purposes and the interpretation of the constants as constants of nature, we are forced to take serious exception to it . . . " [48].

The use of the logistic as an empirical curve to describe population growth implies that the parameters average the factors which produced the observations. When the curve is extrapolated for predictive purposes, this implication is absurd unless the growth of population actually has an inherent stability. At first thought, it may seem preposterous to assume that there can be any inherent stability in population growth. But observations on many human populations in the past reveal that there has been amazing stability over long periods of time in population growth.

Furthermore in forecasting we are interested, from the statistical point of view, primarily in the trend line. For any given year in the future, the observed point may fall above or below the projected trend line; except in unusual instances, it will not fall exactly on the curve. But granting all this, it seems unlikely that the logistic will be used extensively for forecasting population growth in the United States.

I have already pointed out that Logistic I projected from 1910 has predicted the subsequent population with remarkable accuracy. However, Logistic II projected from 1940 probably will run well below the growth of population for at least 20 to 30 years and possibly longer. In other words, it will depart systematically from the probable growth of population. Even though the general trend may be roughly correct, the projected curve is biased. But by what criteria shall we determine that the logistic should be projected from 1900, 1910, 1920 or some other date?

The logistic does not explicitly reveal the relative influence of the elementary demographic factors, births, deaths and migration upon population change. Instead it conceals their separate effects within its parameters. For the analysis of current demographic changes, this is a

serious defect. On the other hand, the work of Lotka has demonstrated that the use of mathematical functions, of which the logistic is one, illuminates many fundamental theoretical problems in demography [49]. Interestingly enough, however, this virtue, as will be pointed out below, is one of the reasons for the failure of certain population forecasts in the past.

There are many statistical difficulties with the logistic such as the fact that a population must be past the point of inflection of the curve before future growth can be described with even rough accuracy and the fact that the curve is not linear in its parameters so that projections of components of the total population present knotty theoretical and practical problems, but these were exhaustively discussed by Wilson and Puffer so that I shall not mention them further here.

There is another, perhaps more important, reason for the failure of the logistic as a predictive curve during the present stage of demographic development of the population of the United States, even though the logistic may be the method of choice for forecasting the future of populations in a different stage of demographic development. This reason I wish to discuss in connection with the explanation for the failure of the true rate of natural increase and the net reproduction rate as methods of forecasting population.

By 1920 some, if not all demographers, realized that the change in the total size of the population was not a reliable guide to the future growth of the population of the United States. As annual statistics of births and deaths became more generally available, population change was analyzed in terms of the crude rate of natural increase and migration.

By 1925, due largely to the work of Lotka, the crude rate of natural increase as a measure of the inherent capacity of a population to replace itself was abandoned in favor of the true rate of natural increase and its correlative, the net reproduction rate. This resulted from the realization that the crude rate of natural increase was affected so strongly by the existing age composition that it was a misleading index of the inherent capacity of the population to replace itself. The net reproduction rate, being independent of the existing age structure, became the measure par excellence, of the inherent replacement power of the population.

A series of annual net reproduction rates were computed. Strangely enough, it did not appear inconsistent to demographers that the inherent replacement capacity of the population should fluctuate from one year to another just as the crude rate of natural increase. No one

seriously questioned the practice of using the fertility and mortality rates of a single year as a measure of long time prospects of population growth; whatever doubts may have arisen were quieted by using the average for several years.

Demographers acted as if hypnotized by the logical power of the theoretical analysis of demographic changes in a stable population and ignored a fact which they knew full well, that the population of the United States was not stable. Actually, except for the temporary age structure of the population, the age-specific fertility rates which are the basis of the net reproduction rate are affected by all the other factors which produce annual fluctuations in the crude birth rate, a rate which had been abandoned as a measure of the inherent capacity for replacement. The sex ratio, variations in the marriage rate both currently and in the past, changes in child spacing, the number of children previously born to a family, response to changes in economic conditions, these and many other factors determine the net reproduction rate of a given year.

As Hajnal so aptly remarked, "the question, 'To what extent is the population replacing itself according to the rates of *this* year?' is a futile question. In however refined a way we analyze the fertility rates of a given year they will still reflect temporary fluctuations" [50]. Whelpton graphically illustrated this fact when he calculated that if the first birth rates derived from the fertility experience of 1942 were to continue unchanged, 1,000 women living through the childbearing ages would produce 1,084 first births, a feat of demographic prowess which undoubtedly would have aroused the envy even of the goddess Diana [51].

Until it was forcefully impressed upon them by the changes in fertility before, during and after World War II, demographers did not appreciate the full implications of a fact which they had been pointing out for several years, the extent to which the number of births during a given interval of time is the result of planned control of fertility. The failure to evaluate correctly the role of voluntary control of fertility operating either through the postponement of marriage or through the deliberate spacing of children after marriage is a major factor in the failure of demographers to foresee with precision the trend of future population growth. The measurement of this factor requires different methods of analysis than those hitherto available or used.

This explanation of the failure of the net reproduction rate to project reliably the future trend of fertility applies equally well to the extrapolation of age-specific fertility rates, for these are but the basic elements of the reproduction rate.

With the widespread adoption of methods of controlling fertility the decision of a couple to have or not to have a child in a given year is strongly influenced by the number of children already in the family and by prevailing economic conditions. This means that not only the number of births but also the average size of existing incompleted families may fluctuate irregularly from year to year. Moreover, family limitation may be exercised before the desired size of family is reached. Thus even though the ultimate number of children is fixed, the spacing of births over the reproductive period may vary widely from time to time.

The only way to form a correct judgment of long time population growth is to analyze carefully the trend in the marriage rate and the number of children per marriage. Except for changes in migration and in mortality, which at the present time exert a relatively minor influence compared with fertility, the principal factor in the long time trend of population growth is the trend in the size of completed families.

Fortunately recent studies indicate that the size of completed families, although declining for at least a century, has showed remarkable stability compared with indexes of current fertility [52, 53]. It may be a matter of doubt as to whether this size ever has dropped below the level necessary to permanently replace the population but if it has, it certainly is not as far below as indicated by the net reproduction rates.

RECAPITULATION

You have heard the charge; I have reviewed the evidence. What is the verdict? I find demographers guilty of

1. Giving the impression that projected populations were relatively inevitable and certain.
2. Underestimating the effects of scientific developments on lowering mortality rates.
3. Believing that the demographic development of a theoretical stable population must inherently characterize current demographic developments in an actual unstable population.
4. Assuming that the observed fertility and mortality of a single year or short period of years is a reliable guide to future trends.
5. Giving undue weight to recent downward cyclical fluctuations in fertility while making long range population projections.
6. Assuming that because birth rates have declined for several generations they must inevitably continue to decline.
7. Forgetting that the voluntary control of fertility can cause the birth rate to rise as well as to fall.

8. Consistently overestimating the rate of decrease in the growth of the population of the United States.

9. Being too uncritical of the work of fellow demographers.

How does this analysis affect the outlook for future population growth? Does it imply that most of the assumptions of demographers concerning the trend of population should be discarded? Does it mean that we should reject, as Professor Davis believes, the view "that growth of the United States population, so rapid as to double in 50–100 years is certainly over"?

Professor Davis' answer to these questions is an unequivocal, categorical yes. "There is now strong reason *not* to expect, within this century, *any* peak followed by levelling off and decline. . . . Barring extreme catastrophes, my own guess is that the population in 2000 will be between 200 and 300 million. But this is only a guess. The range between these projections for the year 2000 is extremely wide; and I believe that no one can yet appraise the factors that will be at work with sufficient accuracy to make a dependable forecast within a range of 50 million. It is often important to realize what we cannot know."

"Furthermore, I make bold to challenge the view—held by almost all demographers of all schools—that our population must, later if not sooner, reach a peak of any size, at any time, from which a decline is probable if not inevitable" [54].

The principal explanation of the fluctuations in the birth rate since 1930 may be found in fluctuations in the marriage rate and in the spacing of children during the reproductive period of family life. As yet there is no conclusive evidence that the size of completed families has increased permanently. However, it is too early to tell whether the high birth rate of the present decade represents merely a "catching up" of postponed births together with some drawing on the future or a real increase in the eventual size of families formed during this period. As of now, the former conclusion appears more likely, although it would not be surprising, if economic conditions remain favorable, for the size of families completed during 1960 to 1965 to be somewhat greater than the size of families being completed at the present time. Any large increase though would represent a major demographic revolution.

In summary, the belief that the rate of population growth in the United States is slowing down remains substantially correct. However, the rate at which it is slowing down and the date on which the total size of the population may become essentially stationary now appear quite different than before the war. The possibility of a decline in total num-

bers before the end of the century seems unlikely, although we should not forget that the birth rate can decline just as rapidly as it can rise.

A preliminary assessment of the possibility of a continuation of the large numerical natural increase of the 1940's through the decade 1950 to 1960 should be possible within the next two years. The total number of registered births during the first 10 months of 1949 was about one per cent larger than the number registered during the corresponding period of 1948. However, the provisional marriage rate based on marriage licenses issued during the first nine months of 1949 is about 14 per cent below the figure for 1948. The number of marriage licenses issued in cities of 100,000 or more, January–October 1949, is lower than the number issued for the same period during any year since 1941 with the exception of 1944. Normally, a drop of this magnitude should be reflected in a drop in the number of births during 1950 and 1951. Whether or not this will be true, only time will tell.

The estimation of population for periods as short as even five years in the future probably will become increasingly difficult unless economic conditions remain stabilized. If the present analysis is correct, the widespread adoption of voluntary methods of birth control make probable a sharp response of fertility to marked fluctuations in economic conditions. This may be true even though the long time trend in population change continues to be well charted. It remains to be seen, however whether a continuation of the present or even a rising standard of living will result in a permanent rise in the average size of completed families. If it does, it will be a new experience in the demographic history of the United States. At present the odds seem to be against such an event happening. It is well to remember, though, that past experience shows future events have little respect for the opinions of demographers and economists alike.

REFERENCES

[1] *Recent Social Trends in the United States.* Report of the President's Research Committee on Social Trends, 1933, Vol. 1, p. xx.

[2] Taeuber, Irene, "Population Studies in the United States," *Population Index*, 12, 254, 1946.

[3] Baker, O. E., "Population Trends in Relation to Land Utilization", *Proc. Second International Conference of Agricultural Economists*, 1930, p. 284.

[4] Thompson, Warren S. and Whelpton, P. K., *Estimates of Future Population of the United States 1940–2000.* National Resources Planning Board, 1943, pp. 3–4.

[5] Bureau of the Census, Series P-5, No. 2, January 31, 1941.

[6] *Statistical Bulletin*, Metropolitan Life Insurance Company, August 1943.

[7] Dorn, Harold F., "The Potential Rate of Increase of the Population of the

United States," *The American Journal of Sociology*, 58, 185–187, 1942.
[8] Whelpton, P. K., "The Meaning of the 1947 Baby Boom." *Vital Statistics, Special Reports*, 33, 7, 1948. (National Office of Vital Statistics.)
[9] *Population Index* 15, 123, 1949.
[10] Davis, Joseph S., "Our Amazing Population Upsurge," *Journal of Farm Economics* (Proceedings Number), November 1949.
[11] Spengler, Joseph J., "Population Predictions in Nineteenth Century America," *American Sociological Review*, 1, 905–921, 1936.
[12] Spengler, Joseph J. *op. cit.*, p. 911.
[13] Bonynge, Francis, *The Future Wealth of America*, 1852.
[14] Gannett, Henry, "Estimates of Future Population," *Report of the National Conservation Commission*, Vol. 2, pp. 7–9, 1909.
[15] Gannett, Henry, *op. cit.*, p. 8.
[16] Pritchett, H. S., "A Formula for Predicting the Population of the United States," *Quarterly Publication American Statistical Association*, 2, 278–286, 1891.
[17] Pritchett, H. S., "The Population of the United States during the Next Ten Decades," *Popular Science Monthly*, 58, 49–53, 1900.
[18] Pearl, Raymond and Reed, Lowell J., "On the Rate of Growth of the Population of the United States since 1790 and its Mathematical Representation," *Proceedings National Academy of Science*, 6, 275–288, 1920.
[19] Verhulst, P. F., "Notice sur la loi que la Population suit dans son accroisissement," *Correspondance Mathématique et physique publiée par A. Quetelet* x, 113–121, 1838.
[20] Lotka, Alfred J., *Elements of Physical Biology*, 1925.
[21] Pearl, Raymond, Reed, Lowell J. and Kish, Joseph F., "The Logistic Curve and the Census Count of 1940," *Science*, 92, 486–488, 1940.
[22] Thompson, Warren S. and Whelpton, P. K., "The Population of the Nation," *Recent Social Trends in the United States*, Vol. I, Chapt. I, p. 46, 1933.
[23] Sharpe, F. R. and Lotka, A. J., "A Problem in Age-Distribution," *Philosophical Magazine*, 21, 435–438, 1911.
[24] Dublin, Louis I. and Lotka, Alfred J., "On the True Rate of Natural Increase," *Journal of the American Statistical Association*, 20, 305–399, 1925.
[25] Böckh, R., *Statistisches Jahrbuch der Stadt Berlin*, 1884, p. 30.
[26] Kuczynski, Robert R., *The Balance of Births and Deaths*, 1928.
[27] Dorn, Harold F., "The Natural Decrease of Population in Certain American Communities," *Journal of the American Statistical Association*, 34, 106–109, 1939.
[28] Dublin, Louis I., "The Outlook for the American Birth Rate," *Problems of Population*, edited by G. H. L. F. Pitt-Rivers, 1932, pp. 115–125.
[29] Whelpton, P. K., *Needed Population Research*, 1938, pp. 6, 7.
[30] Whelpton, P. K., "Population of the United States, 1925 to 1975," *The American Journal of Sociology*, 34, 253–270, 1928.
[31] Whelpton, *ibid.*, pp. 267, 270.
[32] Whelpton, *ibid.*, p. 270.
[33] Whelpton, P. K., "The Future Growth of the Population of the United States," *Problems of Population*, edited by G. H. L. F. Pitt-Rivers, 1932, pp. 77–86.

[34] Whelpton, *ibid.*
[35] *Recent Social Trends in the United States.* Report of the President's Research Committee on Social Trends, 1933, Vol. I, pp. 48–49.
[36] *Recent Social Trends, ibid.*, p. 48.
[37] *The Problems of a Changing Population,* National Resources Committee, 1938.
[38] Whelpton, P. K., "An Empirical Method of Calculating Future Population," *Journal of the American Statistical Association,* 31, 457–473, 1936.
[39] Whelpton, *ibid.*, pp. 470–471.
[40] Bureau of the Census, Series P-3, No. 15, July 23, 1941.
[41] *Estimates of Future Population of the United States 1940–2000,* National Resources Planning Board, 1943.
[42] *Forecasts of the Population of the United States, 1945–1975,* Bureau of the Census, 1947.
[43] *Ibid.*, p. 39.
[44] Bureau of the Census, Series P-46, No. 7, Sept. 15, 1946.
[45] Bureau of the Census, Series P-25, No. 18, Feb. 1949.
[46] *Population Index* 14, 195, 1948.
[47] Davis, Joseph S., *The Population Upsurge in the United States,* War and Peace Pamphlets No. 12, 1949—Food Research Institute, Stanford University.
[48] Wilson, Edwin B. and Puffer, Ruth R., "Least Squares and Laws of Population Growth," *Proc. Am. Acad. Arts and Sciences,* 68, 287–382, 1933.
[49] Lotka, Alfred J., "*Analyse demographique avec application particuliere l'espèce humaine,*" Paris 1939.
[50] Hajnal, J., "The Analysis of Birth Statistics in the Light of the Recent International Recovery of the Birth-Rate," *Population Studies* 1, 137–164, 1947.
[51] Whelpton, P. K., "Reproduction Rates Adjusted for Age, Parity, Fecundity, and Marriage," *Journal of the American Statistical Association,* 41, 501–516, 1946.
[52] Woofter, T. J., "Completed Generation Reproduction Rates," *Human Biology* 19, Sept. 1947.
[53] Whelpton, P. K., "Cohort Analysis of Fertility," *American Sociological Review,* 14, Dec. 1949.
[54] Davis, Joseph S., (ref. 47).

THE POPULATION OF THE WORLD IN THE YEAR 2000

FRANK W. NOTESTEIN
Office of Population Research, Princeton University

The demographic development of the population of the world is divided into the three regions of Incipient Decline, Transitional Growth, and High Growth Potential. For each of these, the factors of changing birth rates, death rates, and family size are considered in their relation to the changes taking place in the social structure and national economy of the countries included in each region.

THIS PAPER is devoted to variations on two themes, neither of which is profound but both of which are too often neglected. The first theme is that the population of the world in the year 2000 will depend very much on events between the present and that date. The second is that the size of the population by the end of the century is much less important than the process by which the population of that date is reached. Neither of these themes denies, of course, that the population of the future depends also on events already past. To come to grips with our problem we must, therefore, first consider the demographic situation and the nature of the processes of change. We can then discuss the ways in which the events of the future may modify the situation. It is only fair to warn the reader that he will not discover here, or elsewhere, how large the population is going to be by the end of the century. The only way I know of finding the answer to that question is to keep alive until that date.

There is little point in discussing the population of the world as a whole. The situation is far too heterogeneous to make it useful to do so, and the problems, although of world-wide importance, are the problems of particular peoples and particular areas. For our purposes it is more useful to consider each of the three types of demographic development first proposed by Warren S. Thompson, for which, however, I prefer the descriptive titles: (1) Incipient Decline, (2) Transitional Growth, and (3) High Growth Potential.[1]

[1] Thompson discussed these types under different names in his book *Plenty of People*, published in 1944. (First edition, Lancaster, Pa., The Jacques Cattell Press; second edition, New York, The Ronald Press Co., 1948.) In fact he used essentially the same types in a paper I had until recently overlooked, which was published in 1929. "Population," *American Journal of Sociology*, XXXIV (6): 959–975. May, 1929.

REGIONS OF INCIPIENT POPULATION DECLINE

The populations in the stage of incipient decline are those that have virtually completed their demographic transition from high fertility and mortality to low fertility and mortality. Their net reproduction rates have been near or below the levels that, if maintained, would ultimately produce a stationary population. Fertility is in large measure under rational control, and mortality prior to the end of the childbearing period is very low. The falling death rates and birth rates of the past have resulted in populations with high average ages. This type includes the populations of Northern, Western, and Central Europe, the United States, Australia, and New Zealand, and more doubtfully those of Canada and Southern Europe. In ordinary course, these populations will grow for a time because their present age distributions tend to support birth rates and to depress death rates. They are properly designated as populations of incipient decline, however, because any perpetuation of the prewar trends in fertility and mortality would lead to decline. In fact they may grow indefinitely, become stationary, or possibly decline. Without immigration, however, they will decline unless the downward trend of the birth rate is permanently checked.

Prior to the war most demographers, and the writer among them, felt that, in the absence of strong positive measures to encourage childbearing, further declines in the birth rates could be expected. There were three main reasons for this expectation. In the first place, the downward trend got under way in a differential fashion, spreading from the upper urban classes of the population down through the social-economic structure and outward from the city to the rural regions. By the end of the interwar period the highest rates were the ones that were declining most rapidly, the lowest were declining least rapidly, and no evidence of a real up-turn was in sight. In the second place, the differences in fertility were closely correlated with differences in the prevalence and effectiveness of contraceptive practice, and there was every indication that contraceptive materials were becoming more abundant, and that the knowledge of their use was spreading. In the third place, the middle class standards, which had been such a strong factor in the motivation for small families, appeared to be spreading throughout the mass of the population. Thus the nature of the trends, the means by which they were brought about, and the nature of the pressures and incentives motivating restrictionist practices all suggested a pattern of decline that had not run its course. On such evidence most of us predicted continued decline unless strong pro-natalist policies were brought to bear.

Instead of further declines, a baby boom got under way with the economic revival in the late 'thirties. The number of births was amazingly sustained during the war, and shot up in the postwar period. We are now in the position of speaking of a secular decline in fertility in spite of a fifteen-year period of generally rising birth rates. Our critics are asking how many years it takes to make a trend, and implying that demographers form an elite corps of double-distilled false prophets—that is, prophets who, being proved wrong, insist on staying that way.

It is evident that the magnitude of the increase in births throughout the Western world has surprised most of us, and that the implications of the rise in births for future population change are not entirely clear. Meanwhile, however, some useful information is at hand. Recent evidence for Great Britain[2] suggests that two technical faults led to some exaggeration of the downward trend of marital fertility during the interwar years.[3] One was the failure, in which I participated, to recognize that the shortage of males caused by the casualties of the First World War and by emigration, was only a temporary depressant of female fertility. The other was the possibility that the swings of depression and recovery, war and its aftermath, may have concealed an essential stability in the size of families of couples marrying since 1927. Our earlier work may therefore have exaggerated somewhat the decline in fertility during the interwar period.

The technical difficulties, however, are relatively minor. Their correction would have resulted only in projections showing slower declines in the birth rates. In fact, there has been a spectacular and unforeseen rise in birth rates throughout the Western world, and the increase has altered the growth characteristics of the population for decades to come. The birth boom has clearly pushed back the date at which a population decline might appear. It is much less certain, however, how large the net addition to births will be when the whole story is in, or that the rise in births represents the launching of an upward trend in family size. All available evidence suggests that the primary source of the additional births was a sharp decline in the age of marriage. When marriage age is dropping, new families are being formed at a faster rate than the population can permanently sustain. After marriage age stabilizes, other things equal, the births will decline. If marriage age should rise the decrease in births would be sharp. The work of the Royal Commission,[4] of Whelpton,[5] and of others, all indicate that the rise

[2] *Report of the Royal Commission on Population.* London, H. M. Stationery Office, 1949.

[3] Hereafter I shall refer to marital fertility, or the average number of children born to married women, as the size of family.

[4] *Super cit.*

came very largely from increases in first and second children. Meanwhile, the specific risks of birth in the very high orders have clearly continued to decline. We cannot be at all certain that there has been any upward trend in family size. Although the possibility of such a trend resulting from early marriage cannot be ruled out, there is no evidence of the re-emergence of the large family. Perhaps the best generalization is that the trend toward decline in family size has been at least temporarily checked.

If the decline in family size has been checked, or perhaps reversed, to what may the check be attributed, and are these factors likely to continue in effect? It seems to me clear that many of the normal peacetime pressures on family size were virtually removed during the war and postwar period. First we had the prewar revival after a deep depression; then, particularly in Europe, the wartime and postwar economic controls; finally in the United States we had the postwar prosperity, and in Europe an extension of child allowance systems. In general money was abundant, goods were scarce, but rationed and subsidized in ways that favored children. In some countries, moreover, military service was less likely for men with children, and maternity and child allowances were common. For the most part governmental policies probably were not intended to sustain births. It may be doubted, however, that there has been a time when child rearing was as heavily subsidized. One also is inclined to wonder whether it will be possible in peacetime to devise an explicit policy for the encouragement of childbearing that is as powerful as the various controls and subsidies of the wartime period proved in fact to be.

One matter in this analysis disturbs me. In retrospect the forces that encouraged births during the war are clear, but this is after-the-fact analysis. Before the fact I, for one, did not think it possible that any policy, however strong, would over-ride the terrors of the wartime experience in Britain. The moral presumably is that having been wrong in the past, I must be prepared to be wrong again, and perhaps am so in the present instance.

So far as the future growth of population of this type is concerned, it seems to me that the following points can be made on the basis of present evidence.

1. The high level of births of the 'forties has definitely postponed the date at which maximum populations could be anticipated in any realistic terms. It is much less clear that the birth boom will have added

[5] Whelpton, P. K., "The Meaning of the 1947 Baby Boom." *Vital Statistics—Special Reports*, 33 (1): 1–10. Oct. 7, 1948.

very much to the population that otherwise might have been expected by the end of the century. Whether it has will depend largely on whether it involves a change in family size or merely a change in the ages at which mothers bear children.

2. In this type of population, in which there is a large measure of rational control of fertility, the annual increments to the population are likely to be irregularly distributed through time because childbearing is sensitive to the swings of the political, social, and economic climates. Such swings produce irregularities in the age distribution that bring complicated problems of social-economic adjustment.

3. Public policy can have a great effect on the birth rate. The prewar analysts recognized this fact, projected future declines without reference to it, and then expressed the belief that an awareness of the trend toward slowing growth would stimulate pro-natalist policies that would modify the trends projected. Because of the war, such policies have come sooner and have been stronger than was anticipated. They have done much more to impair the relevance than the validity of the prewar demographic work.

4. Turning now to prospects for future change, we must first note one contingency that is all too real, yet cannot be usefully taken into account. It is possible that there will be a war of such destructive and disorganizing nature that the growth characteristics of the population will be sharply changed. If so, about the last thing to worry about will be our mistaken estimates of future population. This possibility, therefore, is not taken into account in the present discussion.

Between now and the end of the century, it seems to me that we may expect sharply fluctuating, but rather low, average rates of increase among the populations of this first type. On the one hand, many of the forces tending toward a reduction of family size are likely to continue in effect. On the other hand, we have yet to see a nation approaching a stationary population that did not launch strong measures to stimulate childbearing. I expect that efforts to increase births will be one of the major preoccupations of those concerned with social legislation in the Western world. Indeed, it has already become so in a few nations. Personally, I think it likely that during the next fifty years the growth of this type of population will average less than half of one per cent per year, and that these populations will represent a substantially smaller part of the world's total by the end of the century than they do today. The important fact, however, is not the size, or the proportion, or the speed of growth. The important fact is that populations of this type can check their growth any time it becomes desirable to do so by means

of reducing fertility rather than by that of increasing mortality. I doubt that the provision of an adequate standard of living for the growth that will come in this group presents grave technical or social problems. The greatest danger, it seems to me, is that concern about slowing growth may drive societies to a renewed emphasis on the obligations of the individual to reproduce for the benefit of the state, church, party, or other extra-personal unit. There is danger that the emotional reaction to slowing growth will lead us to seek people for society, rather than to enrich society for the people.

REGIONS OF TRANSITIONAL GROWTH

The next earlier stage of demographic evolution is that of transitional growth. It is given the name because the processes of modernization have already begun to reduce fertility and mortality in ways that are very likely to bring rapid growth. The regions in this stage include Eastern Europe, the Soviet Union, Japan, and parts of Latin America. The populations are heavily dependent on agriculture, but are in the process of urbanization and industrialization. Education and the technological skills are poor but improving. The agrarian family, and the attitudes toward childbearing that accompany it, are still important, but are being modified by the encroachments of an increasingly urban and secularly minded community. Under the influence of such developments, death rates have been declining for a considerable time. The essential characteristic, however, is that even the birth rates, which respond much more slowly than death rates to changes in the social setting, have a well-established downward trend. With death rates leading the birth rates in the decline, there is a substantial margin of growth which may be expected to continue for some time. This stage is one of almost automatic growth, such as that through which Western Europe and the United States were moving in the latter part of the nineteenth and early part of the twentieth centuries. Unquestionably, public policy can shorten or lengthen the transitional period and modify its rate of increase, but it seems unlikely that anything but severe political disorganization could check growth.

Between now and the end of the century, it seems likely that growth will exceed an average of one per cent per year for this population type. In the case of Japan, it seems very doubtful that such growth will, or could, occur, at least in the home islands. Her population, like that of England, is far beyond the size that can be supported by a self-contained economy. The development of foreign markets and sources of food and raw materials before heavy American subsidies are with-

drawn, coupled with emigration and a sharp decline in the birth rate, would seem the only way of avoiding stark tragedy. Otherwise the areas of this type apparently have the resources required to provide better living for more people, as skills and equipment improve. These populations, therefore, represent another sector of the human race that has a reasonably good chance of absorbing its growth long enough to permit the transition to efficient population replacement based on low birth and death rates.

REGIONS OF HIGH GROWTH POTENTIAL

The third of our population types is that of high growth potential. It is the least homogeneous of all, having as its only essential characteristic the fact that the birth rates have thus far given no clear indication of a downward trend. Its death rates in general are high, but some of them are declining rapidly, others little, and it is possible that some are rising. Some of the populations are growing rapidly, others slowly, and, for all we know, some may be declining. The populations are designated as those of high growth potential because, without exception, the birth rates are very high and sufficiently resistant to change to guarantee rapid and sustained population growth any time it becomes possible to achieve something less than extremely high mortality. More than half of the world's population falls into this class, which includes virtually all of the Far East except Japan, most of Africa north of the Union, much of Latin America, and most of the Middle East. Some of the populations are densely settled on the land, others are rather sparsely settled. We may consider the less densely settled regions of this type first.

In much of South America, Borneo, Sumatra, New Guinea, Africa, and the Middle East population densities are low, and unused resources are relatively abundant. If large amounts of capital were available and would be used for development, the regions could carry much larger populations than they do at present, and provide much higher levels of living than exist today. In a period of rapid economic development, the populations would grow very fast indeed, for development would require the control of many diseases now ravaging the people. Moreover, development would cut the death rate long before the slower processes of social reorientation would begin to put pressures on the birth rate. Under these conditions growth might very well exceed two per cent per year. Such growth would not complicate world problems of food supply, however, for if it were to come it would do so because of the expanding productivity of the regions themselves. In one sense the

Middle East is in a particularly favorable position. The presence of oil goes far to solve the immediate problems of available funds with which agriculture can be rehabilitated, industries started, and education, health, and skills improved.

Whether the population of these more sparsely settled regions will grow rapidly is another matter. They will if there is economic development. The difficulties in the future, just as in the past, seem less those of resources than of social, economic, and political organization. Given the solutions to these problems, the processes of population growth should not become major obstacles to an escape from the stable equilibrium of poverty and early death.

It is the people living in the densely settled areas of high growth potential who present the most serious problem. These areas include Egypt, India, China, Korea, Formosa, Java, and most of the Caribbean. The populations are mainly agrarian, have very low levels of living, and are crowded on relatively undeveloped and inadequate resource bases. In these circumstances it is difficult to see how productivity can be increased sufficiently to support rapidly rising levels of living during a period of population growth such as that through which Europe and Japan have moved.

There is no reason to suppose that the average birth rate for the population of this type is as low as 40 per thousand, official reports to the contrary notwithstanding. Nor is there reason to expect that birth rates will decline rapidly. The social organizations, customs, and beliefs governing reproduction remain today those that throughout the centuries have elicited the births necessary for survival in the presence of uncontrolled mortality. These institutions, customs, and beliefs change only gradually, as the experience of the Western world and that of Japan have shown. Unless new elements enter the situation, we may expect birth rates to be little changed for several decades and to remain relatively high during the period with which we are concerned.

Birth rates of 40 per thousand do not guarantee a growing population. They guarantee only the alternative of growth or excessively high death rates. With a birth rate of 40, a population will grow at a substantial rate even if the expectation of life at birth is only one-half that of the present white population of the United States. Something of the potentiality of the situation may be appreciated if we remember that under Japanese rule the population of Formosa increased by 93 per cent in the thirty-five years from 1905 to 1940. Its average annual rate of growth was nearly two per cent. Under Japanese rule, birth rates remained effectively unchanged, but stable government and economic

development, with the protection of health that these entail, stimulated rapid growth. Whenever similar improvements in mortality are possible we may expect similar growth among the populations of this type. Fertility is high enough to maintain increases up to two per cent per year any time death rates can be reduced to levels that the modern nations of the Western world would consider scandalously high for their own people.

If we deal with expectations rather than hopes, I doubt that we should predict rapid population growth at least for the next two decades. The reason for such pessimism—that is the expectation that death rates will be high—lies essentially in the problems of social and political organization that these populations face under difficult economic conditions. The economies must expand substantially just to take care of their increased populations unless death rates are to rise. Resources are not abundant, and capital is very scarce and difficult to accumulate under conditions in which only a little tightening of the belt makes the difference between life and death. Meanwhile, the masses of the population are becoming aware that other people have an easier lot, and are beginning to aspire to higher standards. There is, moreover, a new political consciousness among the people, and new loyalties are replacing old ones. Tight economic conditions, mounting popular aspirations, and divided loyalties make an explosive mixture. There may be a considerable period of disorganization as the populations pass from one regime to another. If so, the danger of catastrophes is very real, as the losses through starvation and epidemics of the past decade show. The margins of life are simply too narrow to permit great retrenchment without mounting death rates. It is by no means impossible that the population growth will be heavily checked by catastrophic losses from time to time and from place to place, and that widespread disorganization could hold growth to an average rate as low as half of one per cent. It should always be remembered that this condition will come only if the world's worst death rates rise spectacularly and hundreds of millions of people die for lack of life's simplest essentials.

There are, of course, other possibilities. One is that strong and wise indigenous governments will provide political stability and, with the assistance of the world, develop agriculture, transportation, and industry at a rapid pace. If this should happen, death rates would drop as a result both of the growing abundance of the means of subsistence and of the public health measures that such changes would require and make possible. It goes without saying that in order to bring about the improved living conditions which such growth implies, the multipliers

for annual economic production would have to be very large. For example, doubled incomes for doubled populations would require a higher multiplier than four in economic production to provide the capital that such economic expansion entails.

Even under relatively favorable conditions, two somewhat different risks seem to me to threaten sustained advances in living levels and health. One risk is that economic production cannot be expanded rapidly enough. Resources are not abundant, and the populations are huge, illiterate, and technologically untrained. Yet, unless economic production can be expanded very rapidly, additions to the population will absorb most of the additions to production. In this case, by the end of the century, our successors will point out, as we, looking at the past, do today that the fruits of development have been mainly more people instead of higher levels of living and health.

The other risk is that by the end of the century the potentiality for further population growth will still be very large. Doubling per capita income for a doubled population in fifty years is one thing. Being prepared to continue the process is quite another. Yet, if we are to judge by the experience of Japan and the Western world the risk is a very real one. It is real if, in the future as in the past, reliance for the reduction of birth rates is placed on the slow and somewhat automatic processes of social change incident to economic development. Birth rates decline under the impact of urban industrial development, rising levels of living, education, and health; but they do so only gradually. Hitherto, the process has involved multipliers for the population of more than threefold, and the circumstances have been much less difficult than those found among the billion people now under consideration. It seems to me that there is grave danger that mounting numbers will be a major obstacle to sustained improvements in health and living levels unless special measures are taken to speed the reduction of the birth rates among the masses of the peasant populations. This task is far from easy. There seems to be no general realization of its importance among the world's leaders. We have little knowledge of the means by which the small family ideal can be efficiently transmitted to agrarian populations, and given the wish for smaller families among such populations, we have at present no sufficiently cheap, effective, and acceptable birth control methods to meet their needs. Between now and the end of the century, the experience of these populations probably will be varied. Some catastrophes are probable, and some successful transitions to efficient replacement are possible. The important fact is that unless growth considerably exceeds one per cent per year, it is unlikely that

living levels will have risen or that birth rates will have fallen very much. If lack of production limits population growth, then little change may be expected either in living levels, or in the social conditions that produce high fertility.

As observed at the outset of this paper, we cannot know what the future population will be, and its size is less important than its modes of change. It seems likely that growth will be rather slow among the populations that have already achieved low mortality and adopted the small family system. The important thing, however, is that these populations can control their growth by modifying their fertility instead of their mortality. Most of the populations in the stage of transition to the small family system may be expected to grow rapidly; but these populations seem likely to achieve their transitions to efficient replacement. Among the sparsely settled populations with high growth potentials, there is room for large expansion. Growth will be rapid if the political and economic obstacles it developed are overcome. It is only in the densely settled populations of high growth potential that growth itself is likely to be a major obstacle to sustained advances in health and material welfare. These populations, however, include more than half of the human race, and their situations are extremely difficult. We do not know that they will grow; we know only that they will grow unless death rates mount, and that growth will be rapid if death rates decline. For them the essential problem is not the size of the population by the end of the century. The crucial problem is whether the intervening processes of social, economic, and political evolution will permit the ultimate termination of growth by means of falling birth rates rather than by means of rising death rates.

Whatever the course of population growth, I doubt that the technological problem of food production will provide the major difficulty. From a technological point of view, man already knows how to produce much more than he is at all likely to produce. The most difficult problems are those in the fields of social science and social engineering. They are the inter-related processes of social, economic, and political change which come to a focus in the problems of population growth. These problems of change should lie at the heart of the social disciplines. Unfortunately, the social sciences, including demography, have little to contribute. We know very little of the processes of change, and are not trying very hard to learn. Yet, it is on the more adequate knowledge of the processes of change in demographic, social, economic, and political fields that the chances for sustained advances in health and material welfare of half the human race may well depend.

POPULATION AND RESOURCES

DISCUSSION OF PAPERS BY FRANK W. NOTESTEIN AND P. V. CARDON

JOSEPH S. DAVIS
Food Research Institute, Stanford University

My general observations, though prepared with reference to both Dr. Cardon's and Dr. Notestein's papers, are all relevant to the latter. First let me summarize a few of their points that deserve emphasis.

(1) The demographic world is not "one world," but a mixed aggregate of many peoples who differ widely in respect of population increase, standards and levels of living, economic potentials, political stability, and other important factors.

(2) As in the past 50 years, rates of population growth in the next 50 will differ strikingly, and change over time, in these various groups of people. We must beware of any form of "globaloney."

(3) The world's economic *potential*, and specifically the potential increase in supplies of food and other wanted goods, is truly enormous, far beyond the ideas of former years; and these potentials can safely be expected to rise.

(4) The extent to which such potentials will be realized in the several population units is wholly unpredictable, since no one can safely forecast how widely or effectively the various procedures will be employed.

(5) The potential increase of population is also enormous, in many specific areas and in the world as a whole.

(6) The size of the world's population in the year 2000 is unpredictable, but this is of no real consequence. As Dr. Notestein clearly implies, it will surely exceed 2,646 million, the upper "asymptote" of Pearl's logistic curve of 1936, unless catastrophes far greater than World War II interfere. But even a figure above 3,000 million should scare no one. What is important is the level of living of whatever numbers there may be.

Next let me add a few points on which Dr. Notestein and Dr. Cardon might agree with me.

(1) The extent to which realized increases in production will (a) be dissipated in defense, warfare, or otherwise, or (b) mainly promote population increase, or (c) be somehow divided between population increase and higher per capita consumption, is unpredictable for the world as a whole. At most, something can tentatively be said about the prospects in some areas.

(2) In the future as in the past, some peoples will enlarge their pro-

ductivity, hold down their numerical increase by preventive checks of various kinds, and achieve levels of per capita consumption increasingly far above "bare subsistence." Others will more nearly conform to Malthus' over-all generalization.

(3) So long as this is true, the disparity between the highest and lowest per capita consumption levels of different peoples will grow wider. Let us not blink this harsh fact. FAO's Director-General Dodd recently termed this disparity a "dangerous gap." If it is dangerous, we must learn to live with it and not expect to eliminate it. I consider wholly untenable, for the next 50 years and even indefinitely, a rash generalization by my late colleague, Carl L. Alsberg, to the effect that the world must ultimately come to a virtually stable population with essentially uniform levels of living.

(4) The important task ahead lies in facilitating improvement in levels of living everywhere, by means that will yield increasingly favorable ratios between population and production wherever they are unfavorable to human progress. The essential elements in this task include not only enhancement of productivity but restraint of population increase where it tends to be excessive. No "Point Four" program, no surplus-distribution schemes, can safely ignore the population factors.

(5) The reasonable prospect lies somewhere between the awful and the millennial extremes that are too frequently presented. Malthus wrote in 1798:

> It has been said, that the great question is now at issue, whether man shall henceforth start forwards with accelerated velocity towards illimitable, and hitherto unconceived improvement; or be condemned to a perpetual oscillation between happiness and misery, and after every effort remain still at an immeasurable distance from the wished-for goal.

Today, the same sort of alternatives are put forward. On the one hand, it is asserted that everyone, everywhere, can, should, and must be well fed, soon, and that "freedom from want" (actually a highly dangerous slogan) is an attainable goal for all the world's peoples. On the other hand, it is argued that population increase and soil depletion seriously threaten calamitous famines within half a century. The history of the past 150 years attests the falsity of Malthus' gloomy prophecy, for the part of the world with which he was chiefly concerned; but it also reveals the illusions of the superoptimists of his day. The history of the next 150 years will likewise presumably confound today's superoptimists and superpessimists.

Finally, let me register dissent from some of Dr. Notestein's views,

chiefly with respect to countries in his first category ("Incipient Decline").

(1) I do not feel that he has fully reckoned with revised appraisals of productive potentials such as Dr. Cardon has set forth, and which H. G. Moulton recently treated in his *Controlling Factors in Economic Development* (Brookings Inst., 1949); or with the prospect that, by and large, the more advanced nations should be in a good position to translate much of this potential into reality in the half century ahead.

(2) He also seems to give inadequate weight to two significant influences in the more advanced nations: (a) prospective gains in life expectancy, and (b) lowered median ages at first marriage and at maternity. If economic conditions improve at a rate well below recognized potentials, both of these gains in level of living bid fair to be maintained. In their effects on population increase, shortening of the time between generations can easily be as important as improved mortality experience.

(3) Dr. Notestein doubts whether the birth boom of the 1940's "will have added very much to the population that otherwise might have been expected by the end of the century." I strongly dissent, at least so far as concerns the United States and Canada. In his Chicago 1944 paper, he considered 176 million a "realistic" forecast for these two countries in 1980 and 2000 (T. W. Schultz, ed., *Food for the World*, Univ. Chicago Press, 1945, p. 53). This figure now looks wholly unrealistic. Their combined population in 1950 will be about 166 million. In view of their abundant resources and relatively high consumption levels, it will be surprising if their net immigration alone does not exceed 10 million in 50 years, though of course it may not add this much to their population in the year 2000. Their combined natural increase will surely be much greater. Neither country suffered heavy war casualties, and mortality experience keeps on improving. Much more important, most of the unprecedented flood of babies in the 1940's will survive to have children before and after 1980. A natural increase of only 20 million in 50 years would imply incredibly low fertility rates. In these two countries, I see no reason to expect that public policies designed to increase births will seem urgent in the next 50 years.

(4) As Dr. Notestein almost realizes, the three neat categories of peoples (which Thompson and he had developed) are unsatisfactory for the analysis of future population growth. If they are not abandoned, subcategories are urgently needed, and the listing of nations requires reconsideration. The Food Research Institute has recently published my study of *The Population Upsurge in the United States* (Decem-

ber 1949). I now feel sure that a new category should be established for at least the United States, Canada, Australia, and New Zealand. If they ever truly fitted in a class labeled "Incipient Decline," they no longer do. Rather, they belong in a new group of countries in which a really substantial rate of population increase can, and probably will, be maintained by a combination of very low mortality, moderate fertility, and moderate immigration. In such countries, developing productivity promises to yield the happy combination of significant rise in levels of living *and* unexpectedly if not absolutely high rates of population increase. Normally, and I think probably, this group of countries in the year 2000 will show a higher percentage of the world's population than it did in 1940.

CONTROLLED SELECTION—A TECHNIQUE IN PROBABILITY SAMPLING*

ROE GOODMAN AND LESLIE KISH

A sampling technique is defined as introducing control into the selection of n out of N sampling units when it increases the probabilities of selection for preferred combinations of units (and decreases the probabilities for non-preferred combinations). Methods used in the past have by no means exhausted the possibilities of controlled selection, however. Procedures are developed by which the probabilities of selection for preferred combinations are sharply increased and the theoretical basis for the methods is stated. The methods are applied to a specific problem and the procedures are described in detail. It it found that as a result the variances of estimates for several important items are reduced as compared with the corresponding variances for stratified random sampling.

AMONG THE decisions involved in designing a sample is that of the method of selecting the sampling units. Simply stated the problem is the following: Given a population consisting of N rigorously defined sampling units, how is one to select n of these units for the sample?

In probability sampling the essential condition is that each of the N sampling units shall have a specific, known probability of selection and that the probability for none of the units shall be zero. Within this limitation the possibilities are myriad.

The methods of controlled selection presented in this paper appear to have their major (although by no means their only) usefulness in the selection of first-stage sampling units in multi-stage sampling.[1] For the sampling of widely dispersed populations first-stage units commonly consist of counties, part-counties, or combinations of them (areas containing perhaps 1000 to 5000 square miles). Although the techniques of controlled selection have wide, general applicability in all types of probability sampling, it was for the selection of first-stage units that these methods were developed and the emphasis in this paper is upon their application to this particular problem.

STRATIFIED SAMPLING

A selection procedure which is widely used today is called "stratified random." By this method sampling units are selected at random within

* Presented at the 108th Annual Meeting of the American Statistical Association, Cleveland, December 29, 1948.

[1] Many of the terms used in this paper are as defined in Statistical Paper, Series C No. 1 of the United Nations, "The Preparation of Sampling Survey Reports," Lake Success, N. Y., Jan. 1, 1949.

each stratum independently of the selection of units within other strata. During recent years units within a stratum have frequently been assigned varying probabilities[2] of selection. When this is done the strata are usually defined in such a way that only one unit is to be selected from each.

STRATIFICATION VIEWED AS A METHOD OF CONTROL

It is to be noted that stratification introduces restrictions or controls in the process of selection. In simple random sampling the probability of selection of any of the $C(N, n)$ possible combinations of n out of N sampling units is a function solely of the probabilities of selection assigned each of the n units. With stratified random sampling on the other hand the selection of units is partially controlled in that the probability of selection of the combination depends on the strata with which the n units are associated. For example, if two units are to be selected from each stratum, every combination of n units consisting of two units from each stratum has a positive probability of selection, whereas all other combinations have a zero probability of selection. By means of stratification then, the number of possible sample combinations is reduced. At the same time the probabilities of selection of admitted combinations are increased.

CONTROLLED SELECTION—DEFINITION

As we define it, the expression "Controlled Selection" has a rather broad meaning. This expression is defined to mean any process of selection in which, while maintaining the assigned probability for each unit, the probabilities of selection for some or all preferred combinations of n out of N units are larger than in stratified random sampling (and correspondingly the probabilities of selection for at least some non-preferred combinations are smaller than in stratified random sampling).

It is to be noted that while in stratified random sampling the probabilities of selection for certain combinations are increased, the possibilities of increasing the probabilities of selection for preferred combinations are by no means exhausted by this process. That is, with con-

---

[2] Although the word random is generally taken to connote equal probability, its use in this paper implies merely specific, known probabilities, not necessarily equal. The notion of making a selection among units with varying assigned probabilities was introduced about six years ago. See Hansen and Hurwitz "On the Theory of Sampling from Finite Populations," *Annals of Mathematical Statistics*, XIV: pp. 333–362, 1943, and "On the Determination of Optimum Probabilities in Sampling," *Annals of Mathematical Statistics*, XX: pp. 426–432, 1949.

trolled selection it is in general possible to go very much further in increasing the probabilities of selection of preferred combinations of units than is done by stratified sampling alone. At the same time the probability of selection for vast numbers of other combinations is reduced, generally to zero. It is with the use of controls after the possibilities of stratification have been exhausted that this paper is concerned.

NEED FOR CONTROLS BEYOND STRATIFICATION

In selecting first-stage units for a sample survey of a large geographical area it is usually desirable to stratify in respect to a number of classifications of the population. The need for this multiple stratification arises in part because various kinds of information may be obtained from the same survey; furthermore a sample of first-stage units may need to service over a period of time a great number of diverse surveys with a multitude of variables as objectives often unknown in advance. At the same time the number of strata one can use is restricted to the number of first-stage units which in turn is limited by considerations of cost. Moreover, the use of additional modes of stratification is possible only at the expense of reducing the control, hence possibly the efficiency, of other classifications. Even with a few classifications one may find himself with more strata cells than the number of first-stage units he can use.

The sampler may now ask this question: Is it possible to achieve controls, similar to those of additional stratification without actually increasing the number of cells? Several types of controls have been designed and used in the past; there is reference to these in the latter part of this paper. These procedures aimed at obtaining balance with respect to control factors, but usually accepted the possibility of biased estimates. In contrast, the emphasis in this paper is upon securing unbiased estimates through rigorous procedures of probability sampling, at the same time only approximating a balance with respect to the control factors.

OBJECTIVE OF CONTROLLED SELECTION

As for most sampling techniques the objective of controlled selection is to reduce the variances of estimates obtained at a given cost. To the extent that the choice among alternative methods of selection does not affect costs appreciably the objective is simply one of reducing the variances. As will be shown later, by extending the control beyond that of stratification it is possible either to reduce the variance of any item or to increase it depending upon the utility of the controls used.

Since it was logical to suppose that the procedures contemplated would result in some reduction in the variances of estimates a sample was drawn for the North Central States by the new method. Later additional test samples were drawn from the same population and the between-first-stage-unit component of the variance was computed for several important items. It was found that these between-components were reduced from 11% to 32% below those corresponding to stratified random sampling. The details of these tests are given later in this paper.

SIMPLE ILLUSTRATION OF CONTROL BEYOND STRATIFICATION

It is possible to introduce a measure of additional control in the selection process very easily. Suppose we have two strata consisting of first-stage units from each of which one unit is to be selected. Suppose further that in stratum 1 units *B*, *C*, and *F* lie adjacent to the ocean or other major waterway, that in stratum 2 unit *d* is similarly located, and that all other units are located inland. Assume the units, and the probability of selection assigned to each of them are as shown below:

Stratum 1		Stratum 2	
Unit	Probability	Unit	Probability
A	.10	*a*	.15
B	.15	*b*	.30
C	.10	*c*	.10
D	.20	*d*	.20
E	.25	*e*	.25
F	.20		

The sum of probabilities for *B*, *C*, *F* and *d* is 0.65. It is considered desirable to select one inland and one coastal unit, and undesirable to select two coastal units. The procedure is as follows:

Re-arrange the units in the first stratum by listing *B*, *C*, *F* first, followed by *A*, *D*, and *E*. Re-arrange the units in the second stratum by shifting *d* to the end, that is, by placing *e* above *d*. Then draw a random number from 1 to 100 and let this one number determine the selection in both strata. Thus if the random number is 45 or less a coastal unit will be selected in statum 1 and an inland unit in stratum 2, if the number is 46 to 80 an inland unit will be selected in both strata, and if the number is greater than 80 an inland unit will be selected in stratum 1 and a coastal unit in stratum 2.

By means of the above procedure the probability of selecting one

inland and one coastal unit is .65 and the probability of selecting two inland units is .35. With stratified random sampling on the other hand the probabilities are the following: one inland and one coastal unit, .47; two inland units, .44; two coastal units, .09. If one were more interested in securing what appeared to be a balanced sample than a strictly random sample he would be inclined to make sure that he selected one inland unit and one coastal unit to represent the two strata. By a procedure of controlled selection, however, the original assigned probabilities of each unit are rigorously maintained and the probability of securing one inland and one coastal unit is made as large as possible within the limitation of probability sampling.

APPLICATION TO A SPECIFIC PROBLEM

There follows now a description of the application of controlled selection to a larger problem. As will be seen, the procedure is somewhat complex, although its concurrence with the rules of probability sampling is apparent. The description may also be helpful in showing the logical basis for the use of controlled selection. In giving the details of the procedure that was used there is no pretense of showing the best way of controlling the sample selection but merely one way that it may be done.

The objective was the selection of 21 primary sampling units to represent the North Central States. The Detroit, Cleveland, Chicago, and St. Louis metropolitan areas were taken as separate strata, each of which was then selected with certainty. There remained the selection of 17 primary sampling units to represent all areas in the region outside these metropolitan areas. The first-stage sampling unit consisted of a single county or a county joined with part(s) or all of an adjacent county (or counties). In forming the primary sampling units, there were 3 objectives as follows:

(1) To increase diversity within the unit with regard to concentration of population.

(2) To limit the distance from the central city in which the interviewer would live, to the periphery, to a range of 25 to 30 miles.

(3) Not to increase the size of the unit beyond a single county if doing so would decrease the proportion of the population residing in the central city to less than 40% of the total for the unit. The last objective was a concession to interviewers many of whom prefer to do a substantial part of their work near their homes.

The sampling units were grouped into 17 strata of about 1.8 million each, primarily on the basis of the size of largest city, and secondarily,

for the more rural counties, by major type of farming areas. The strata were numbered from 1 to 17, the ordering being from those with largest cities to those with rural populations only. Within each stratum one unit was to be selected; the probability of selection assigned each unit was proportional to its 1940 population.

It was thought possible, by the use of controls, to insure a greater geographic spread and balance than one could expect from simple stratified random sampling. Another aim was to achieve a better balance with regard to per cent urbanization in the middle 7 strata, in which there remained considerable variation within the strata with respect to this variable. The geographic balance was intended to yield approximately proportionate representation of the various states in the sample and to make sure that the larger cities, as well as the less densely populated places, were well distributed geographically.

SELECTION IN SUCCESSIVE STAGES—STEP ONE

In order to simplify the mechanics of the work it was found desirable to divide the procedure into two successive steps. The various schemes used to accomplish control in these two stages will be discussed with the aid of the tables.

The first step was designed to assure proper representation among four groups of states as shown in the heading of Table I. The table shows the sums of the probabilities of selection assigned the different sampling units within each of the four groups (*A*, *B*, *C*, and *D*) for each of the 17 strata, and for various groups of strata 1 to 4, 5 to 8, etc. These probabilities are proportional to the population figures of the 1940 Census and expressed as a fraction of the total population for the stratum. The ordering of the strata is meaningful, representing decreasing size of the central city within the sampling units. For example stratum 1 consists of three densely populated units, stratum 6 of thirteen units having medium sized towns, and stratum 17 is made up of 195 entirely rural units.

The data in the table may be interpreted in terms of expected number of selections, as well as probabilities of selection. For example, the total 1.1718 under *B* for strata 5 to 8 means that of the 4 units to be selected in these strata, on the average 1.1718 will be selected in Ohio or Indiana. By the use of controlled selections it is possible *to make sure that in any one sample either one or two of the four units selected will be drawn from this pair of states*. By this method, when a draw is made the chances are .1718 that two units will be selected from these states and .8282 that a single unit will be selected. As will be seen, this is the major type of control achieved by the first step.

TABLE I

ASSIGNED PROBABILITIES OF SAMPLING UNITS CLASSED IN FOUR STATE GROUPS IN EACH OF 17 STRATA, NORTH CENTRAL STATES

Strata of approximately equal total populations, assigned probabilities adding to one. Strata order by decreasing size of cities.

Stratum No.	Mich., Wis. A	Ind., Ohio B	Ill., Iowa, Minn. C	Mo., N. D., S. D., Neb., Kan. D	Total
1			.4946	.5054	1.0000
2	.4146	.5854			1.0000
3		1.0000			1.0000
4	.2827	.3296	.2602	.1275	1.0000
Total 1–4	.6973	1.9150	.7548	.6329	4.0000
5	.1059	.1093	.5165	.2683	1.0000
6	.4937	.3213	.1850		1.0000
7	.1034	.7412	.1554		1.0000
8	.4361		.3478	.2161	1.0000
Total 5–8	1.1391	1.1718	1.2047	.4844	4.0000
Total 1–8	1.8364	3.0868	1.9595	1.1173	8.0000
9	.3040	.4550	.1601	.0809	1.0000
10		.2222	.4737	.3041	1.0000
11	.2915	.0759	.2981	.3345	1.0000
12	.0146	.2476	.4657	.2721	1.0000
13	.2666	.2149	.1922	.3263	1.0000
Total 9–13	.8767	1.2156	1.5898	1.3179	5.0000
14	.0261	.2719	.4297	.2723	1.0000
15	.2914	.1891	.2963	.2232	1.0000
16	.0363	.0663	.3409	.5565	1.0000
17	.2164	.0299	.1570	.5967	1.0000
Total 14–17	.5702	.5572	1.2239	1.6487	4.0000
Total 9–17	1.4469	1.7728	2.8137	2.9666	9.0000
Total 1–17	3.2833	4.8596	4.7732	4.0839	17.0000

In order to show how the procedure was carried through we turn to a discussion of Table II. This will be followed by a description of one procedure for obtaining such a table.

Table II is set up in such a way as to conform with Table I. The number of cells in each box has been limited to 16 in order to simplify the work. The numbers in the cells represent the number of sampling units to be drawn from each group of states. The P's are the probabilities of selection for each combination; they add to 1.0000. Thus the first combination (represented by the first box) has a probability of

TABLE II

THIRTEEN ALTERNATIVE COMBINATIONS WHICH JOINTLY SATISFY THE REQUIREMENTS OF TABLE I

P=Probability of selection for the combinations; A, B, C, D denote state groups. Entries show the number of units to be selected in each cell.

Strata	No. of Strata	$P=.1391$				$P=.0033$				$P=.0835$				$P=.0334$				$P=.0434$				$P=.0850$			
		A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
1–4	4	0	2	1	1	0	2	1	1	0	2	1	1	0	2	1	1	0	2	1	1	1	1	1	1
5–8	4	2	1	1	0	1	2	1	0	1	2	1	0	1	1	1	1	1	1	1	1	1	2	1	0
9–13	5	1	1	2	1	1	1	2	1	1	1	1	2	1	1	2	1	1	2	1	1	0	2	2	1
14–17	4	0	1	1	2	1	0	1	2	1	0	2	1	1	1	1	1	1	0	1	2	1	0	1	2
Total	17	3	5	5	4	3	5	5	4	3	5	5	4	3	5	5	4	3	5	4	5	3	5	5	4

Strata	$P=.2047$				$P=.0022$				$P=.0383$				$P=.0838$				$P=.1404$				$P=.0940$				$P=.0489$			
1–4	1	2	0	1	1	2	0	1	1	2	0	1	1	2	1	0	1	2	1	0	1	2	1	0	1	2	1	0
5–8	1	1	2	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9–13	1	1	2	1	1	1	2	1	0	2	2	1	1	1	2	1	1	1	1	2	1	1	1	2	1	2	1	1
14–17	0	1	1	2	0	1	1	2	1	0	1	2	0	1	1	2	1	0	2	1	1	1	1	1	1	0	1	2
Total	3	5	5	4	3	5	4	5	3	5	4	5	3	5	5	4	4	4	5	4	4	5	4	4	4	5	4	4

.1391 of selection. If this combination were selected, the drawings for the first four strata would be taken as follows: none from the A group, two from the B group, 1 from the C group and 1 from the D group. The number of units to be selected from each group is similarly shown for strata 5–8, 9–13, and 14–17.

The probabilities of all cells of Table I are preserved in Table II, that is, the product in Table II of the entry in a cell and the probability for the combination when summed over all combinations equals the entry for the corresponding cell in Table I. For example for group A in strata 1–4 we have: $(0)(.1391)+(0)(.0033)+\cdots+(1)(.0940)+(1)(.0489)=(0)(.3027)+(1)(.6973)$. In other words: the probability is .3027 of selecting no A, and .6973 of selecting one A for strata 1–4. Again for strata 5–8 the probability is .8609 of choosing one A, and .1391 of choosing 2 A's. Again for strata 1–8 combined we have for group A: $(1)(.1636)+(2)(.8364)=1.8364$. Similarly for the other cells and finally for the total of all strata 1–17 the probability is .7167 of selecting three A's, and .2833 of selecting four A's. The same restrictions hold for groups B, C and D over all the totals shown in Table I. Also for all the totals of Table I, the corresponding entry, or sum of entries, of Table II never differs by more than a fraction.

There are thirteen combinations shown in Table II; this appeared to

be the least necessary to satisfy the requirements of the seven lines of totals of Table I.

There are various devices by which a set of combinations such as those in Table II may be derived. One manner of arriving at this result is described below. It should be noted that the solution given in Table II is not unique. Rather, when one faces two possible ways of completing a combination, one may select that which appears preferable. In the present example this freedom was utilized to control the number of units coming from groups A and B combined, hence also from C and D combined. Contrarily, if a desired restriction cannot be entirely satisfied, because of conflicting restrictions one may nevertheless work toward minimizing the occurrence of an undesirable combination.

OBTAINING CONTROL TABLE II

We begin by writing down a desirable combination such as the one in the first box in Table II. To do this, we place a zero on the first line under A, since from Table I it is found that the probability for this cell is only .6973. In other words, the number in this cell must sometimes be zero. (We could have just as well written a first combination in which this entry was 1, it doesn't matter.) When the first entry is zero the second should be 2, since the sum of the entries in Table I (.6973+1.9105) is greater than 2, hence the sum of the first two entries should never be less than 2 (since a control was attempted in respect to $A+B$ and $C+D$). Having made these entries, the other two entries on the first line should each be 1, since the total must be 4 (the number of selections from the first four strata) and neither entry should be greater than 1 since the probabilities for the cells (.7548 and .6329) are each less than 1. On the second line we place a 2 under A, since the entry for the cell in Table I (1.1391) is greater than 1 and the total for the first 8 strata under A (1.8364) means that the sum of the first 2 entries under A should usually be 2. Having entered the 2 under A, the entries of 1, 1, 0 under B, C, D, respectively, automatically follow since the sum of the four entries must be 4 and neither of the entries under B or C should be less than 1, since the probabilities in Table I (1.1718 and 1.2047) are both greater than 1. The entries on the third line are those in the corresponding cells in Table I rounded to whole numbers. They add to the required total of 5. The total for the last 9 strata in Table I (1.4469) indicates that the sum of the last two entries under A should more often be 1 than 2. A zero is accordingly entered on the last line under A. The entry under B should then be 1, since the sum of entries in Table I (.5702+.5572) is slightly greater than 1. The entry of 1 under C and 2

under D results in totals of 3 for the last two lines in each of these columns, which totals are in harmony with the corresponding totals of 2.8137 and 2.9666 in Table I. It may be noted also that the totals for the entire columns in this combination (3, 5, 5, 4) are in harmony with totals in Table I. The probability of .1391 is the largest desirable probability for this combination, since the entry on the second line under A can be a 2 only this proportion of the time, if it is never to be less than 1, and the assigned probabilities are not to be violated.

Table III is useful at this point in that it shows the permissive probabilities for each desired number of selections in each cell. The .1391 assigned the combination just discussed is subtracted from each permissive probability corresponding to the cell numbers of selection appearing in this particular combination. As additional combinations are set down the probabilities are successively subtracted until the remainders are all zero. Thus we are guided in assigning probabilities to the different combinations in that the restrictions imposed by Table I are made ironclad by means of Table III. For example, it is obvious from the table that the maximum probability for the first combination is .1391, since the subtractions in the first column leave a zero on line A, strata 5–8. The second and third combinations were originally written as one, that is, the second combination was assigned the probability of .0868; however, it was later found convenient to split this combination and slightly alter the arrangement in the lower right corner yielding the third combination. The probability of .0868 was the maximum for a combination having a total of 4 for the first eight strata under B.

It is perhaps fairly easy to see how certain "improvements" may be made in the combinations and their probabilities as entered in Table II. For example, the combinations could probably have been written in such a way as to harmonize with totals for strata 5 to 13 in Table II, as well as with the other totals. Also it is readily seen that there was freedom to use a much greater number of combinations than were used, if so desired. The procedure that was used, however, was based on the belief that until it is established that techniques of this kind result in gains in sampling precision, efforts may well be concentrated on the more obviously desirable aspects.

Table II is intermediate to Table IV, which provides a basis for selection of specific groups of sampling units. Using the data in Tables I and II, a table such as Table IV can quite readily be written down. It will be noted that for each box in Table II there is a corresponding column (or perhaps a set of columns) in Table IV. In setting forth the data in Table IV it is possible to do more than to conform with the re-

TABLE III

WORK SHEET SHOWING RESIDUAL PROBABILITIES FOR EACH DESIGNATED NUMBER OF SELECTIONS

Cell designation		Selections in each cell		Probabilities of combinations to be successively deducted												
Strata	Group	Number	Probability	.1391	.0033	.0835	.0334	.0434	.0850	.2047	.0022	.0383	.0838	.1404	.0940	.0489
1– 4	A	0	.3027	.1636	.1603	.0768	.0434	.0000								
		1	.6973						.6123	.4076	.4054	.3671	.2833	.1429	.0489	.0000
	B	1	.0850						.0000							
		2	.9150	.7759	.7726	.6891	.6557	.6123		.4076	.4054	.3671	.2833	.1429	.0489	.0000
	C	0	.2452							.0405	.0383	.0000				
		1	.7548	.6157	.6124	.5289	.4955	.4521	.3671				.2833	.1429	.0489	.0000
	D	0	.3671										.2833	.1429	.0489	.0000
		1	.6329	.4938	.4905	.4070	.3736	.3302	.2452	.0405	.0383	.0000				
5– 8	A	1	.8609		.8576	.7741	.7407	.6973	.6123	.4076	.4054	.3671	.2833	.1429	.0489	.0000
		2	.1391	.0000												
	B	1	.8282	.6891			.6557	.6123		.4076	.4054	.3671	.2833	.1429	.0489	.0000
		2	.1718		.1685	.0850			.0000							
	C	1	.7953	.6562	.6529	.5694	.5360	.4926	.4076		.4054	.3671	.2833	.1429	.0489	.0000
		2	.2047							.0000						
	D	0	.5156	.3765	.3732	.2897			.2047	.0000						
		1	.4844				.4510	.4076			.4054	.3671	.2833	.1429	.0489	.0000
1– 8	A	1	.1636		.1603	.0768	.0434	.0000								
		2	.8364	.6973					.6122	.4076	.4054	.3671	.2833	.1429	.0489	.0000
	B	3	.9132	.7741			.7407	.6972	.6123	.4076	.4054	.3671	.2833	.1429	.0489	.0000
		4	.0868		.0835	.0000										
	C	1	.0405								.0383	.0000				
		2	.9595	.8204	.8171	.7336	.7002	.6568	.5718	.3671			.2833	.1429	.0489	.0000
	D	1	.8827	.7436	.7403	.6568			.5718	.3671			.2833	.1429	.0489	.0000
		2	.1173				.0839	.0405			.0383	.0000				
9–13	A	0	.1233						.0383			.0000				
		1	.8767	.7376	.7343	.6508	.6174	.5740		.3693	.3671		.2833	.1429	.0489	.0000
	B	1	.7844	.6453	.6420	.5585	.5251			.3204	.3152		.2344	.0940	.0000	
		2	.2156					.1722	.0872			.0489				.0000
	C	1	.4102			.3267		.2833						.1429	.0489	.0000
		2	.5898	.4507	.4474		.4140		.3290	.1243	.1221	.0838	.0000			
	D	1	.6821	.5430	.5397		.5063	.4629	.3779	.1732	.1710	.1327	.0489			.0000
		2	.3179			.2344								.0940	.0000	

TABLE III—*Continued*

Cell designation		Selections in each cell		Probabilities of combinations to be successively deducted												
Strata	Group	Number	Probability	.1391	.0033	.0835	.0334	.0434	.0850	.2047	.0022	.0383	.0838	.1404	.0940	.0489
14–17	A	0	.4298	.2907						.0860	.0838		.0000			
		1	.5702		.5669	.4834	.4500	.4066	.3216			.2833		.1429	.0489	.0000
	B	0	.4428		.4395	.3560		.3126	.2276			.1893		.0489		.0000
		1	.5572	.4181			.3847			.1800	.1778		.0940		.0000	
	C	1	.7761	.6370	.6337		.6003	.5569	.4719	.2672	.2650	.2267	.1429		.0489	.0000
		2	.2239			.1404								.0000		
	D	1	.3513			.2678	.2344							.0940	.0000	
		2	.6487	.5096	.5063			.4629	.3779	.1732	.1710	.1327	.0489			.0000
9–17	A	1	.5531	.4140					.3290	.1243	.1221	.0838	.0000			
		2	.4469		.4436	.3601	.3267	.2833						.1429	.0489	.0000
	B	1	.2272		.2239	.1404								.0000		
		2	.7728	.6337			.6003	.5569	.4719	.2672	.2650	.2267	.1429		.0489	.0000
	C	2	.1863					.1429							.0489	.0000
		3	.8137	.6746	.6713	.5878	.5544		.4694	.2647	.2625	.2242	.1404	.0000		
	D	2	.0334				.0000									
		3	.9666	.8275	.8242	.7407		.6973	.6123	.4076	.4054	.3671	.2833	.1429	.0489	.0000
1–17	A	3	.7167	.5776	.5743	.4908	.4574	.4140	.3290	.1243	.1221	.0838	.0000			
		4	.2833											.1429	.0489	.0000
	B	4	.1404											.0000		
		5	.8596	.7205	.7172	.6337	.6003	.5569	.4719	.2672	.2650	.2267	.1429		.0489	.0000
	C	4	.2268					.1834			.1812	.1429			.0489	.0000
		5	.7732	.6341	.6308	.5473	.5139		.4289	.2242			.1404	.0000		
	D	4	.9161	.7770	.7737	.6902	.6568		.5718	.3671			.2833	.1429	.0489	.0000
		5	.0839					.0435			.0383	.0000				

TABLE IV

SELECTION PATTERNS SHOWING THE STATE GROUPS FROM WHICH THE VARIOUS SAMPLING UNITS WILL BE SELECTED

P = probability of selection of the pattern
P cum = cumulated probabilities of patterns

P	.0200	.0061	.0038	.0195	.0124	.0657	.0116	.0002	.0031	...	.0489
P cum.	.0200	.0261	.0299	.0494	.0618	.1275	.1391	.0002	.0033	...	.0489
P cum.							.1391		.1424	...	1.0000
Stratum No.											
1	C	C	C	C	C	C	D	D	D	...	C
2	B	B	B	B	B	B	B	B	B		A
3	B	B	B	B	B	B	B	B	B		B
4	D	D	D	D	D	D	C	C	C		B
5	A	A	A	A	A	C	A	B	B		D
6	C	C	C	C	C	A	C	C	C		B
7	B	B	B	B	B	B	B	B	B		C
8	A	A	A	A	A	A	A	A	A		A
9	B	C	C	C	C	C	C	C	D		B
10	C	B	B	B	B	B	B	D	B		D
11	D	C	C	C	C	C	C	A	A		A
12	C	D	D	D	D	D	D	C	C		C
13	A	A	A	A	A	A	A	B	B		B
14	D	D	C	C	D	D	D	D	C		D
15	C	C	D	D	B	B	B	C	A		C
16	D	D	D	B	C	C	C	D	D		D
17	B	B	B	D	D	D	D	A	C	...	A

The table means that the pattern in the first column has a probability of .0200 of being chosen, the second one a probability of .0061, and so on. As indicated by the dots, many of the patterns are omitted here for lack of space. If the first pattern were selected the sampling unit for the first stratum would be selected from those units in state group C, the sampling units for the second and third strata in state group B, the sampling unit for the fourth stratum in state group D, and so on.

The patterns have been written down in such a way as to conform with Tables I and II. The sum of the probabilities of selection of the patterns in which a certain letter appears on a given line (stratum) is equal to the probability shown for that letter and stratum in Table I. For example, for C in stratum 1: $.0200+.0061+.0038+\cdots+.0489=.4946$. The patterns are grouped in sets, each set corresponding to a combination in Table II, as indicated by the cumulated probabilities on the second line of the above table. Within each set within each group of strata (1–4, 5–8, etc.) each letter will occur the number of times required for that combination in Table II. For example, in each pattern of the first set ($P=.1391$) in strata 1 to 4, there are zero A's, two B's, one C and one D. Moreover in writing down the patterns an attempt was made to avoid writing the same letter in two adjacent strata in the same pattern. Thus in each pattern the different letters (state groups) are well distributed.

Each of these patterns of course would have a chance of selection if ordinary stratified sampling were used. However the sum of the probabilities of all of them *under ordinary stratified sampling would be only .0000058 in contrast to the 1.0000 here.*

strictions of Tables I and II, namely, to see that the selections in any group of states are scattered more or less uniformly throughout the strata. For example, in the first column, in which there are to be two selections from group *A* in strata 5 to 8, the arrangement is such that these are not taken from numerically adjacent strata. The same holds for the two selections from group *C* in strata 9 to 13, and the two selections from group *D* in strata 14 to 17. The two selections from group *B* in strata 1 to 4, on the other hand, are adjacent, due to the fact that all of stratum 3 is in group *B*; hence, whenever the selection for stratum 2 is taken from group *B*, the two *B*'s must appear in successive strata.

In writing down the selection patterns, as in Table IV, the manner in which Table I is used may be illustrated in the following way: According to Table II, one selection in the first 4 strata is to be taken from units in group *C* in all but three of the combinations. Table I shows that the letter *C* must be entered opposite stratum 1 in selection patterns having a combined probability of .4946 and opposite stratum 4 in selection patterns having a combined probability of .2602. Entries for each letter can therefore be made one at a time for the first four strata until all of the patterns are completed as far as these strata are concerned. A similar procedure is then used for the other sets of strata.

In view of the small number of units in each of the first six strata (there are only three units each in the first two) a still more exacting system of controls seemed feasible for these strata. Here the process was extended to the setting down of combinations of individual sampling units with a probability of selection determined for each such combination. In this way, it was possible to emphasize desirable combinations of sampling units, that is, those yielding a good geographical scattering of the larger cities.

At this point a random selection was made between .0001 and 1.0000. Number .8184 was drawn, and if Table IV were given in full the selected pattern could be found there, namely, the first one in which the cumulated probability equals or exceeds .8184.

One may rest at this point if he wishes, assuming that he has obtained the substantial part of whatever gains controlled selection may yield. He would then make his selection of one unit for each stratum at random among the sampling units of each selected group. That is what was done for purposes of the calculation of variances.

SECOND STAGE OF CONTROLLED SELECTION

On the other hand one may introduce further restrictions within the restrictions of the groups selected from Table IV. In this instance, it

was decided to secure an approximate balance in respect to distribution among individual states and also in respect to per cent urbanization. At this stage this can only be done, however, by applying a procedure of controlled selection to the groups selected in step one. All the units in other cells of Table I are henceforth out of the picture altogether. Thus, the balancing of the ultimate sample in respect to these additional factors cannot be better on the average than that of the groups already selected.

A table, somewhat comparable with Table IV was prepared for the second stage of sample selection. It set forth the possible selection patterns for selection of sub-groups and again was not a unique system. Prior to the preparation of the table, sub-groups of units within the selected groups were formed on the basis of state, and within state, on the basis of whether the units were plus, average, or minus, in respect to per cent urbanization. For these purposes the probabilities of selection for each sub-group were scaled upward in such a way that the total probability for each group was 1.0000. Frequently, the units in a given sub-group belonged to only one of the states in the group; hence, the selection from the group could be drawn only from this one state. It would have been possible, of course, to have obtained a greater control of the sample distribution by states by having done additional work prior to the first selection. Moreover, at this, the second stage, it would be possible to introduce a measure of geographic control within the states if this were considered worth-while.

After a random draw was taken, based on this last table, there were eleven sets of sub-groups, some of which contained but two or three sampling units. (The individual sampling units for the first six strata had been automatically determined by the procedure of step one.) Additional random draws were taken within sub-groups until, finally, the selection of seventeen sampling units was completed.

COMPARISON OF VARIANCES

To obtain estimates of the variances for the controlled selection 100 samples of 17 units each were drawn using Table IV. Within the selected groups individual units were chosen with probabilities proportionate to their 1940 populations, as required by the original sample design. This means that the samples were selected at the conclusion of step one of the control procedure; whatever further gains step two may have yielded were not measured since the selection of 100 samples by this method would have entailed considerably more work. The mean of the 17 units for each sample was obtained and then the variance

among the 100 means. The variance of the stratified random selection was computed by means of the standard formula which utilizes the information for the entire population.

For the purposes of this comparison items were selected mainly from 1940 Census data. Items were sought for which data were published for each township (as well as for counties) because some of the sampling units have boundaries which cut across county areas along township lines. Another item included was the 1944 Presidential vote because of interest in this variable; in the states for which the vote was not published by townships, the vote for split counties was estimated by splitting it in proportion to the population.

Upon completion of the computations it was apparent that the items studied fell into two categories with respect to the magnitudes of the variances. Since in practice the selected first-stage sampling units are in turn subsampled, a "within" component of variance must be added to the between variance in order to secure a total variance comparable to those of estimates derived from an actual survey. For this purpose the within variance was computed using the formula appropriate for a simple random sample of n cases within each of the 17 first-stage units and assuming that the variance within each of these units was the same. When the between variances were compared with "total" variances it was found that the items studied fell into the two following classes: (1) those for which the between variance was an important part of the total variance and (2) those for which the between variance was a minor part of the total variance.

The results of the computations are presented in Table V. In order to show the relative importance of the between component for the different items the between variance for stratified random sampling is also given as a percent of the total variance for cases of $n=25$ and $n=100$.

From this table it is evident that for items 1, 2, 3, and 4 the between component is an important part of the total variance for n's as small as 25 within the primary sampling units; while for items 5, 6, 7 and 8 the between component is relatively unimportant even for n's as large as 100. Now the practical value of a procedure designed to reduce the between component of the variance must be judged by its effect on those items for which the between component is important. For items for which the between component is small even a drastic reduction of the latter will have no great effect on the total variance. In Table V this means that the control procedure must be judged by its effects upon items 1, 2, 3 and 4.

Results of the computations showed that the use of controlled se-

TABLE V

"BETWEEN" COMPONENT OF VARIANCES DERIVED FROM CONTROLLED SELECTION AND FROM STRATIFIED RANDOM SAMPLING

(17 First-Stage Units in North Central States)

Item No.	Item	Population mean	Stratified random			Controlled selection	
			Between variance	Between variance as per cent of total variance		Between variance	Between variance as per cent of stratified random between variance
				$n=25$	$n=100$		
1	Average monthly rent ($)	21.62	1.015	.63	.87	0.785	77
2	Per cent dwellings vacant	5.70	1.129	.47	.78	0.879	78
3	Per cent foreign born	5.71	0.934	.42	.75	0.638	68
4	Per cent voting Democratic	45.98	3.896	.40	.73	3.452	89
5	Dwelling units per 100 persons	29.10	0.363	.07	.23	0.356	98
6	Per cent 65 years and over	8.20	0.125	.07	.22	0.080	64
7	Per cent adult females	31.83	0.165	.03	.11	0.146	88
8	Per cent 5 years and under	8.00	0.039	.02	.08	0.037	94

In items 3, 6, 7, and 8 the categories are expressed as percentage of the total population (1940 Census).

The average monthly rent is from the 1940 Census, except that the within p.s.u. variance is unknown. A rough estimate of $16 for the standard deviation was used. The vacant dwelling units are expressed as per cent of total dwelling units (1940 Census).

Item 4 represents the votes cast for the Democratic Presidential candidate in 1944 as a per cent of the votes cast for all candidates, as given in official state manuals.

Item 5 represents the number of dwelling units per 100 persons in the population (both from the 1940 Census).

lection resulted in reductions of 23, 22, 32, and 11, per cents respectively in the between variances for items 1, 2, 3, and 4. In terms of the total variances the reductions are 14, 10, 13, and 5 per cents respectively for n's of 25, and 20, 17, 24, and 8 for n's of 100.

Gains of this kind clearly justify the additional work involved in carrying through this procedure. From one viewpoint the reduction of 20 per cent in the between variance is roughly equivalent to having spread a given number of interviews over a number of primary sampling units one-quarter greater than originally contemplated. Of primary importance of course is the fact that for a given degree of accuracy a smaller sample is possible, insofar as the item is one for which the between component is important.

It is of interest to note that for items 5, 6, 7, and 8, for which the be-

tween component was unimportant, the reduction in variance was small, but for each of these items the procedure did result in a slight gain. Theoretically as will be shown, this procedure can result in an increased variance; and it is to be expected that for occasional items in which the between variance is small it will result in increases. Even so, of course, such increases could be expected to be of slight importance in comparison with the total variances for such items.

SOME PREVIOUS INVESTIGATIONS OF SELECTION TECHNIQUES

Earlier research workers seem to have dealt with more limited applications of controlled selection. There is, of course, a wide range of published material on stratification, including a fundamental contribution by Neyman[3] in 1934. Neyman also analyzed the usual type of purposive selection which he found to be generally inferior to stratified random sampling.

Previously Bowley[4] and Jensen[5] had reported their analyses of stratified random sampling and purposive selection but their findings had been far from conclusive. Jensen,[6] in 1928, described the purposive selection of a sample from records of the 1923 Danish Agricultural Census and showed that it represented the population well in respect to distributions of several farm variables. Strand and Jessen[7] compared the use of purposive and stratified random selection of townships in Iowa counties and concluded that "purposive selection does not provide samples of greater accuracy than stratified random selection" for situations of the type investigated. However, none of these investigators attempted to combine purposive selection with probability sampling.

Frankel and Stock,[8] in 1942, suggested the use of a Latin square design which, within the framework of the modes of stratification used, extended the controls of stratification to a second dimension. The following year Tepping, Hurwitz, and Deming[9] reported extensive analy-

---

[3] Neyman, J., "On the Two Different Aspects of the Representative Method," *Journal of the Royal Statistical Society*, Vol. 97, 1934, pp. 558–625.

[4] Bowley, A. L., "Measurement of the Precision Attained in Sampling," *Bulletin de l'Institut International de Statistique*, Tome XXII, 1926.

[5] Jensen, A., "The Representative Method in Practice," *Bulletin de l'Institut International de Statistique*, Tome XXII, 1926, pp. 381–439.

[6] Jensen, A., "Purposive Selection," *Journal of the Royal Statistical Society*, Vol. 91, 1928, pp. 541–547.

[7] Strand, Norman V. and Jessen, Raymond J., "Some Investigations of the Suitability of the Township as a Unit for Sampling Iowa Agriculture," Iowa State College Research Bulletin 315, 1943.

[8] Frankel, Lester R. and Stock, J. Stevens, "On the Sample Survey of Unemployment," *Journal of the American Statistical Association*, Vol. 37, 1942, pp. 77–80.

[9] Tepping, Benjamin J., Hurwitz, William N , and Deming, W. Edwards, "On the Efficiency of Deep Stratification in Block Sampling," *Journal of the American Statistical Association*, Vol. 38, 1943, pp. 93–100.

ses on techniques of this kind which they designated as "deep stratification." There is an important limitation of a Latin square design, namely that the probability of selection of each of the $L \times L$ cells into which the population units are grouped is the same. Tepping, Hurwitz, and Deming accordingly considered estimates, derived as though the probabilities for the various cells were equal, whether or not the combined probabilities for the units in the different cells were equal in all cases. That is, they considered the use of biased, as well as unbiased, estimates and in some cases found the bias to be an important source of error. Yates[10], in 1946, reported a selection procedure called balancing in which additional random selections are substituted for units originally drawn until finally "the mean value of the balanced factor in the sample is equal to the mean of the factor in the whole population." He stated that estimates derived from such a sample are "subject to some elements of bias" but concluded that "such bias is not likely to be of importance in practice."

CONTROLLED SELECTION AS AN EXTENSION OF PURPOSIVE SELECTION

Conceptually, the use of controls in selecting a sample may be viewed as an extension of the technique known as purposive selection (although perhaps involving more judgment). If, however, the estimates to be derived from the sample are to be unbiased, an additional step not ordinarily considered to be a part of purposive selection is required. In order that the sampling may be probability sampling, the sampler must select not just one but many purposive samples, until every unit in the universe is included in one or more samples. The number of samples in which each unit appears must be exactly proportionate to its assigned probability of selection.

After the complete set of purposive samples has been established, the random selection of one of them constitutes a probability sample. As has been seen, such a procedure is not in conflict with the use of stratification but, on the contrary, can be more readily accomplished after strata have been established.

The preceding describes what appears to be the ultimate in controlled selection. In the process of purposive selection one could, conceivably, use considerable judgment and also make numerous checks in regard to various known characteristics of the sampling units, finally establishing samples each of which was as nearly as possible in accord with the population as a whole with respect to each characteristic.

[10] Yates, F., "A Review on Recent Statistical Developments in Sampling and Sampling Surveys," *Journal of the Royal Statistical Society*, Vol. 109, 1946, pp. 12–43.

Whether a procedure of this intricacy would ever be operationally feasible though is very questionable.

CONTROLLED SELECTION VIEWED AS SYSTEMATIC SAMPLING

Despite certain similarities of controlled selection to deep stratification and balancing there appears to be a still greater resemblance between this procedure and a method which has been increasingly discussed in the literature in the last few years, namely systematic sampling. Madow and Madow[11] and Cochran[12] have shown the theoretical conditions under which systematic sampling can be expected to be useful and have discussed the application of the method to certain real situations. In the research on systematic sampling to date, the emphasis has been upon the selection of a sample systematically from a more or less "naturally ordered" sequence of sampling units. Very recently, however, Madow[12a] has discussed the notion of selecting from the different strata in such a manner as to induce a negative intra-class correlation but his paper attacked the problem solely from a theoretical point of view.

The similarity between the highly controlled selection discussed above and systematic selection is apparent when it is considered that if the N units in the population are numbered in the proper sequence the selection of a sample systematically will automatically yield one of the purposive samples. In other words, under this condition the two methods become the same.

Systematic sampling, in fact, appears to be a good common denominator for various methods of probability selection. Unrestricted random selection may be viewed as systematic selection in which the sampling units in the universe have been ordered at random.[13] Assuming a uniform sampling rate within all strata, stratified random selection is systematic selection in which the strata are placed in some sequence and the sampling units within each strata are randomly ordered.[13] The

---

[11] Madow, William G., and Madow, Lillian, "On the Theory of Systematic Sampling, I," *Annals of Mathematical Statistics*, Vol. XV, 1944, pp. 1–24.

[12] Cochran, W. G., "Relative Accuracy of Systematic and Stratified Random Samples for a Certain Class of Populations," *Annals of Mathematical Statistics*, Vol. XVII, 1946, pp. 164–177.

[12a] Madow, William G., "On the Theory of Systematic Sampling II," *Annals of Mathematical Statistics*, Vol. XX, 1949, pp 333–354.

[13] In the selection of units with varying probabilities the analogy to systematic sampling holds only when exactly one unit is to be selected from each group. This is, however, the usual manner in which varying probabilities are utilized in practice. In such instances, one may proceed as follows: Express the probabilities assigned the various units in fractions having a common denominator, d. The sum of the numerators of these fractions is, of course, d. Then arrange the d chances of selection in a random order. It is well to recognise that with this procedure the various chances of selection relating to a single unit may be scattered throughout the sequence.

selection with deep stratification, assuming equal probabilities for all cells, becomes a systematic one in which the cells within each stratum are ordered in a randomly chosen sequence to accord with the restriction of a Latin square and the sampling units within each cell are also randomly ordered.[13] Thus controlled selection may be viewed as systematic selection in which the ordering of the sampling units within strata has in part at least been done purposively.

Viewed in this way the vast amount of flexibility inherent in methods of controlled selection is evident. Moreover, the theory of systematic selection can be expected to provide valuable clues regarding the conditions under which various procedures of controlled selection may be useful.

THE UNDERLYING THEORY

As the Madows[14] have stated, systematic sampling may be viewed as cluster sampling, in which each possible sample is one of the clusters. If the expected intra-class correlation for a cluster of n units is negative,[15] less than $-1/(N-1)$ to be precise, the estimated mean has a smaller variance than that of an unrestricted random sample of n units. Moreover the variance decreases with decreasing values (increasing negatively) of the intra-class correlation.

In (proportionate) stratified random sampling the intra-class correlation is never positive, and usually it is less than $-1/(N-1)$. But the Madows and Cochran have pointed out that under certain conditions a systematic sample will have a smaller variance than a stratified random sample (which is a systematic sample with the units ordered at random within strata); that is, that the intra-class correlation for the former will be less than that for the latter.

Now, unless the intra-class correlations corresponding to all of the possible arrangements of units within the strata are identical, the values corresponding to some of the arrangements must be less than the average of all of them, which is that of stratified random sampling. It follows that in practice there always exist orderings of the units which have a smaller intra-class correlation (and therefore smaller variance) than that of stratified random sampling. Equally inescapable is the fact

---

[14] *Op. cit.*

[15] As indicated in Footnote 13, units with varying probabilities may be viewed as having varying numbers of chances of selection, the sum of which was given as d. The problem of selection is then restated as that of selecting n units from a population in which the sum of the chances of selection for all units is N, in which case the number of units as such in the population becomes irrelevant. Thus, the conclusions to be reached apply alike to selection with equal and varying probabilities subject to the restriction that when varying probabilities are used one and only one unit is to be selected from each stratum or sub-group.

that the ordering may increase the intra-class correlation above that of stratified random sampling.

Many questions are still to be answered. There is the problem of how to obtain optimum ordering in respect to a variable for which information is sought. Then there is the need to resolve the probable conflict between optimum orderings for two or more variables to be investigated in the same survey. For example, the optimum ordering for estimating one variable may result in less accuracy than would stratified random sampling as far as another variable is concerned. The question becomes one of finding an ordering of the units by means of which the intra-class correlations will be reduced on the average.

It is also pertinent to consider the possible reductions of the variances for different sizes of sample. The Madows found that with a given ordering of the units, the effect of change in sample size was somewhat complex and that further investigation of this problem was necessary.

In summary, the theory shows that the use of additional controls may reduce the variance of estimates for some items in a survey but also that the result may be an increased variance for these or other items. It remains then for empirical studies and additional theoretical developments to determine what the results are likely to be in practice.

CONCLUDING REMARKS

The emphasis in this paper has been upon the possibilities for using controlled procedures of selection without violating principles of probability sampling. Although, conceivably it may appear to be desirable to reduce the probabilities for all "non-preferred" combinations to zero even at the risk of violating the assigned probabilities there is here contained no evidence on this question. The procedures outlined do eliminate the possibility of selection of vast numbers of combinations but among the millions remaining it may indeed be difficult to determine that any particular combinations will necessarily comprise inferior samples. At this writing there seems to be little justification in accepting less than complete objectivity in the final selection of first-stage units.

The procedures described illustrate a selection with uniform rates within strata of approximately equal size. The methods are sufficiently flexible, though, to permit their ready extension to disproportionate sampling with strata of different sizes.

Multi-stage sampling (the entire process) may itself be viewed as controlled selection in which the selection of first-stage units is simply

the first phase in the process of selecting the ultimate physical units. In such a controlled selection the preferred combinations as a rule have larger variances than for stratified random sampling but the increased variance may be more than offset by reduced cost. Thus the preferred combinations of controlled selection are not necessarily the ones yielding the smallest variances because there are frequently cost differentials to be considered.

There are many applications of controlled selection in the later stages of multi-stage sampling and elsewhere. For example in selecting a sample of towns within selected first-stage units it may be worth-while to secure an approximate balance over the entire sample with respect to such characteristics as size of place, and average rental value and perhaps within individual first-stage units deliberately to secure an "unbalance" with respect to geographical location (this last in order to reduce travel from town to town within a county).

A problem that will require attention is the estimation of variances from sample data when controlled selection has been used. This is the problem of estimating variances for systematic sampling and while it presents difficulties it appears possible to develop approximations which may be so formulated as to be on the "safe side."

The results of the tests reported in this paper indicate that it is possible to obtain appreciable gains of precision through the use of the procedures of controlled selection proposed here. If the gains obtained in these tests are found to be typical for large scale surveys generally, the efficiency of sample surveys based on the relatively small number of sample areas can be increased through the use of these procedures. Additional tests are needed in relation to different variables, and also in relation to the use of ratio and other more complicated estimates.

A TECHNIQUE FOR ANALYZING SOME FACTORS AFFECTING THE INCIDENCE OF SYPHILIS*

B. G. Greenberg, Ph.D.; John J. Wright, M.D., M.P.H.; and Cecil G. Sheps, M.D., M.P.H.

Institute of Statistics and School of Public Health, University of North Carolina

In interpreting rates of syphilis incidence, it is worthwhile to evaluate certain factors which influence the probability of acquiring syphilis. In the past, these factors have been estimated separately either by graphical or by elementary numerical methods. A model is proposed in the present paper to measure these factors simultaneously. The parameters in the model can be estimated by multiple regression and tested for statistical significance.

I. INTRODUCTION

In most countries where systematic reporting of disease is practiced, incidence rates of syphilis are periodically observed in populations specific for race and sex, and classified by age. The present investigation proposes a method of interpreting these rates in terms of some of the important factors which affect an individual's probability of acquiring syphilis.

Two factors which are immediately apparent are the age distribution of the new cases and the percentage change in the rates from year to year. These factors have been treated to some extent in studies [3], [10], [11], [12], and [13] and will be discussed later in more detail. Vonderlehr and Usilton [25], using the method of life tables, based upon observed cases in 1936–1937, obtained an exact age distribution of the probability of acquiring syphilis. Another type of life table, based upon observations obtained by following throughout life a group of individuals born at approximately the same time in order to measure their chance of acquiring syphilis at each age, suggested the inclusion of a third factor. That factor was the effect of the date of birth upon a person's chances of acquiring syphilis.

The circumstances tying these three factors together can be expanded, at first, to include mortality and morbidity rates in general. Suppose a series of such rates for a specific race and sex group have

* The statistical investigation was made possible by the grant of an assistantship from the Institute of Statistics. The syphilis studies are supported jointly by grants from the International Health Division of the Rockefeller Foundation and the North Carolina State Board of Health.

The authors wish to express their appreciation to Professor H. L. Lucas, Jr., and Professor W. G. Cochran for advice on special problems.

been observed over a period of time and presented in a two-way classification table. One variable in the classification is time, and the second variable is usually age. If we let r_{ij} represent a rate of this kind at age a_i $(i=1, 2, \cdots, m)$ occurring at time t_j $(j=1, 2, \cdots, n)$, the data may be arranged in the following manner:

Age	Time						Average During Period
	t_1	t_2	t_3	$\cdots$	t_{n-1}	t_n	
a_1	r_{11}	r_{12}	r_{13}	$\cdots$	r_{1n-1}	r_{1n}	$\bar{r}_1$
a_2	r_{21}	r_{22}	r_{23}	$\cdots$	r_{2n-1}	r_{2n}	$\bar{r}_2$
a_3	r_{31}	r_{32}	r_{33}	$\cdots$	r_{3n-1}	r_{3n}	$\bar{r}_3$
$\cdot$	$\cdot$						
$\cdot$	$\cdot$						
$\cdot$	$\cdot$						$\cdot$
a_{m-1}	r_{m-11}	r_{m-12}	r_{m-13}	$\cdots$	r_{m-1n-1}	r_{m-1n}	$\bar{r}_{m-1}$
a_m	r_{m1}	r_{m2}	r_{m3}	$\cdots$	r_{mn-1}	r_{mn}	$\bar{r}_m$
All Ages	$r_{.1}$	$r_{.2}$	$r_{.3}$	$\cdots$	$r_{.n-1}$	$r_{.n}$	

Without any loss in generality, assume that both factors, time and age, are equally spaced in yearly intervals. Let α_i represent the true rate associated with age a_i, and τ_j the true effect due to time t_j. It is customary to measure the effect of time $(\tau_1, \tau_2, \cdots, \tau_n)$ as percentage changes in the mortality or morbidity under question by comparing $r_{.1}, r_{.2}, \cdots, r_{.n}$. This procedure is followed, for example, in computing standardized or age-adjusted rates. Similarly, the usual practice for measuring the rates associated with age $(\alpha_1, \alpha_2, \cdots, \alpha_m)$ is to average in some fashion the rates over a sample period of time in order to obtain $\bar{r}_1, \bar{r}_2, \cdots, \bar{r}_m$.

The reasoning behind such practices is the belief, whether so stated or not, that underlying the rates, r_{ij}, there exists a fairly stable age distribution of occurrence which combines with time in such a manner as to produce the observed results. It seems likely, however, that a rate r_{ij} may not be completely described by combining only the effect of age (α_i) with the effect of time (τ_j). That is, an interaction factor between age and time may be necessary in the model to describe other influences. For example, an epidemic at time t_j may affect one particular age group, whereas another epidemic at time $t_{j'}$ may attack a different age group.

An illustration of what is meant by this interaction will be helpful. Suppose the probability of a male between ages 10 and 11 dying within

one year was 0.02 during the last decade. Further, assume that the probability of dying for all males in all age groups had increased 10% from 1940 to 1942 because of the intervening war. It is obvious that males, age 10, would have a probability of dying during 1942 less than $0.02+.10(0.02)=0.022$. The reason is that such males constituted a cohort of persons born in a common span of time who would not have been greatly affected by the cause for the change in the probability of death from 1940 to 1942.

By introducing into the model a factor representing the interaction between age and time, (mn) additional parameters are created, resulting in more constants than observation points. Restrictions on the parameters must be imposed if the scheme is to be of practical value. A useful set of restrictions on the interaction factors is suggested by considering the cohort of persons born in a common span of time. On the table of the two-way classification, the diagonal lines enclosing r_{11}, r_{22}, r_{33}, $\cdots$, represent one such cohort. The restriction used is that all the interactions between age and time along this diagonal line are equal. Thus, the cohort measures the effect of date of birth upon the mortality or morbidity in question.

The notion of interaction in the sense of a cohort of persons born during a common era was used successfully by Frost [8] in studying age changes in tuberculosis mortality. His purpose was to suggest that the age distribution of tuberculosis mortality not be studied from the column $\bar{r}_1, \bar{r}_2, \cdots, \bar{r}_m$, but rather from a group of persons born at the same time, such as the cohort represented by rates $r_{11}, r_{22}, r_{33}, \cdots$. This method achieved one of the objectives of a long term follow-up study. Frost pointed out graphically that by creating several cohorts and tracing the tuberculosis mortality of these individuals throughout life, the shape of the age distribution was remarkably constant.

This paper may be considered an attempt to generalize and quantify the principle of cohorts. A method is proposed for measuring the time changes in the various cohorts while estimating simultaneously the effects α_i and τ_j.

In an analysis with cohorts, such as is suggested, there are $(n+m-1)$ possible cohorts. Each cohort has a distinct number of observations, distributed according to the following plan, where s is the smaller of (n, m):

Chronological Rank of Cohort	1	2	3	$\cdots p \cdots q \cdots$	$n+m-3$	$n+m-2$	$n+m-1$
Number of Observations in Cohort	1	2	3	$\cdots s \cdots s \cdots$	3	2	1

There are h values of s, where $h=g-s+1$ and g is the greater of (n, m). According to Frost's technique, only the h or $(h+2)$ central cohorts were worth graphing because of the paucity of observations at the two extremes. In the present analysis, it is suggested that solely for the purpose of measuring cohort effects, it is worthwhile to reduce, by grouping, the $(n+m-1)$ possible cohorts to a smaller convenient number. That is, the data themselves are not pooled, but the effects due to the date of birth are condensed. This blocking of cohort effects will be considered again when the problem of syphilis incidence is discussed.

In order to illustrate the method and to demonstrate the practical nature of the results, data for an example were selected from a study whose reliability was known to the authors. Before proceeding to the main part of the paper, a brief description of the data and the manner in which the rates of syphilis incidence were derived are presented in the following section.

II. THE DATA ON RATES OF SYPHILIS INCIDENCE

The data for this example are from the North Carolina Syphilis Studies which have been described elsewhere [28], [29]. The study area comprised the Orange-Person-Chatham Health District and the Durham City-County Health District of North Carolina. The former represents a rural area involving three counties, whereas the latter is predominantly urban. The population in each health district was approximately 75,000 in 1940.

Infectious syphilis constituted the new cases and was defined for the purpose of the North Carolina Syphilis Studies as primary and secondary syphilis. Although syphilis is a reportable disease, it is recognized that only infectious cases which were discovered and diagnosed were reported to the health authorities. All reported cases were administered supervised treatment either by their own physicians or by local health authorities. Every effort was made to discover all cases occurring in the study area. There were no material changes in the methods of discovery in the study area once the research program started. Every attempt was made to obtain complete case histories. All records were cross-checked to eliminate duplication and error.

The data for Negro females were selected for this example to assure that the number of cases reported would be adequate for statistical analysis. Few females were absent from home because of military duty, and a larger share of the syphilis problem is associated with the Negro population. The Negro population constituted about one-third of the

total population. The analysis was performed by considering the four counties as a unit without any attempt to point out urbanization differences.

A further restriction on the data involved the limiting of the example to the ages 15–29 inclusive. Although infectious cases among Negro females occur outside this range in a scattered fashion, little doubt exists that the problem of infectious syphilis is largely confined to this age group.

The number of cases of syphilis are for the years 1941–1947. Since they must be considered in conjunction with respective exposed popu-

TABLE 1

AGE SPECIFIC DISCOVERY RATES OF PRIMARY AND SECONDARY SYPHILIS AMONG NEGRO FEMALES PER 1000 POPULATION, 1941-1947, BY AGE AND YEAR OF DISCOVERY IN STUDY AREA

Age*	1941	1942	1943	1944	1945	1946	1947
15	1.6	5.0	10.2	8.5	1.7	24.4	8.2
16	9.4	12.9	16.6	15.2	15.5	22.6	6.4
17	14.0	16.0	21.5	16.8	20.5	27.1	4.7
18	10.3	15.1	26.8	36.2	23.5	34.8	17.9
19	20.6	19.6	15.7	23.0	16.8	30.0	13.3
20	7.4	18.0	21.8	32.8	20.8	42.0	15.5
21	15.1	4.9	11.1	26.7	21.0	20.2	18.7
22	5.5	10.4	15.3	29.4	7.8	17.2	11.0
23	7.0	5.7	7.3	21.7	8.0	21.2	6.8
24	0	7.3	2.0	11.6	7.4	9.1	3.5
25	6.7	1.7	7.3	1.9	1.8	8.0	5.1
26	5.1	1.8	7.4	1.9	2.1	10.6	5.0
27	3.7	3.5	3.7	3.8	4.0	5.8	5.2
28	1.8	1.9	0	5.7	0	1.8	3.8
29	1.8	5.5	1.9	0	3.8	12.4	0

* The ages reported were to nearest birthday. The ages in this table, therefore, represent the midpoints of yearly intervals.

lations to obtain attack rates, there was a need for postcensal estimates of the population on a county basis classified according to sex, race and age.

Ordinarily, one would utilize a census tract or population register in this regard [5]. There was none available for any portion of the study area. At the time of writing, no state agency in North Carolina made a county population estimate independently from the Bureau of the Census. The latter has sporadically furnished estimates of population data, its characteristics and changes. A description of the postcensal population data available is summarized in [14]. With these as a basis,

crude estimates of the Negro female population in the four counties were made. The method is explained in detail in Appendix 1.

In computing the age specific attack rates, there arose the question of "selection without replacement." That is, if an appropriate rate indicating the probability of acquiring syphilis is to be determined, the

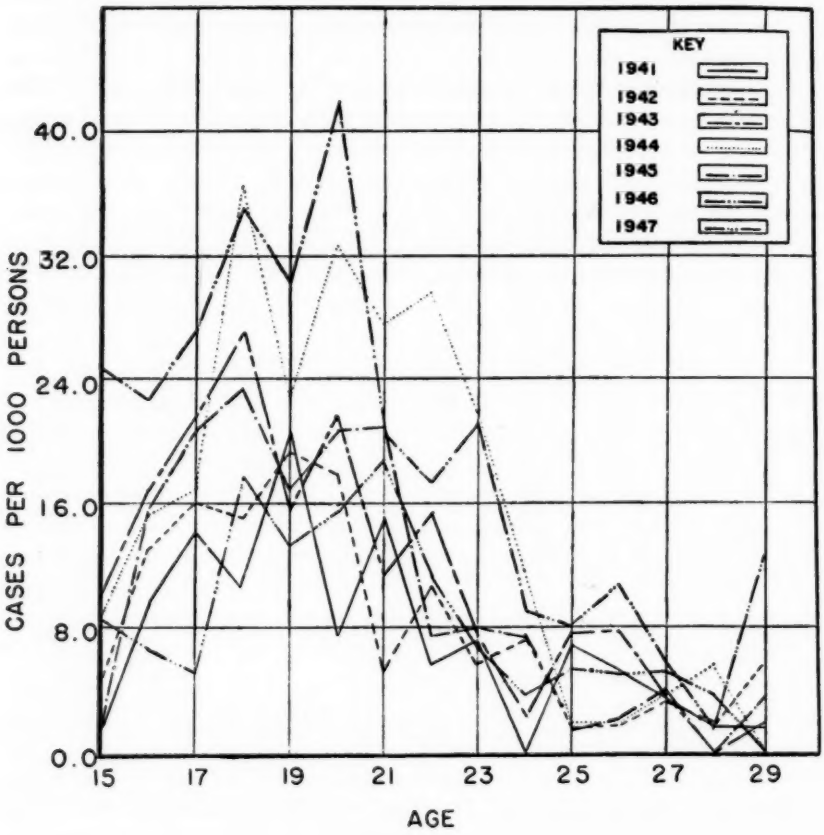


FIGURE 1

AGE SPECIFIC DISCOVERY RATES OF SYPHILIS INCIDENCE FOR NEGRO FEMALES IN STUDY AREA, 1941-1947, BY AGE

average population of susceptibles is necessary. Therefore, at the time an individual is diagnosed as having infectious syphilis, adjustment should be made in the total population at all succeeding dates to render it a population of susceptibles. The infected person should be considered non-susceptible until there is assurance that she can be infected again

with syphilis. For purposes of this study, it was assumed that all treated cases were susceptible to re-infection, since the average discrepancy in this example was less than 2% even if such an assumption was completely wrong.[1] The annual rates of syphilis incidence are presented in Table 1 and Figure 1.

III. FORMULATION OF MATHEMATICAL MODEL

The problem here was to decide upon the appropriate model relating the rate of incidence to age, year and date of birth. The model was constructed to be as simple as the nature of the data permitted. Previous experience and various theoretical considerations were called upon as aids in formulating the model.

Instead of fitting an individual parameter for each of the m age groups, it is not unreasonable to postulate that there is a basic frequency distribution of new cases dependent upon age. To estimate the shape of this curve, it can be seen from Table 4 of Vonderlehr and Usilton [26] that acquired syphilis is first observed in early adolescence, its incidence reaches a peak soon thereafter, and then diminishes rapidly. Such a relationship may be represented by a Pearson Type III curve ($y = c_1 X^{c_2} e^{c_3 X}$). There are possibly other related distributions which might define this relationship a little more accurately but which are not as simple. Support that the Pearson Type III curve is appropriate is given by examination of the graphs of the data in Figure 1.

It was assumed that the age distribution is basic and that the year and date of birth effects are superimposed. That is, the expected

[1] A common tool employed by actuaries to make automatically a portion of the adjustment necessary in this example is the use of central rates [2], [24]. Under this method, the population at the middle of the period is adjusted by one-half the number of cases during the period. That is, if p_x is the age specific probability of being infected with syphilis at age x, then

$$p_x = \frac{I_x}{M_x - (1/2)I_x}$$

where I_x = number of infectious cases of syphilis occurring during period, and
M_x = number of persons in population at middle of period.

This type of central rate presupposes, however, that an attack confers absolute immunity from re-infection. There is experimental evidence to the contrary, although it is usually impracticable to distinguish between relapse and re-infection. Furthermore, if p_x is based upon annual periods, this central rate fails to consider any residue of immunity among individuals infected in previous years with a correction for those who died. Under present circumstances, it is impossible to estimate the degree of immunity (if any) which is conferred and to estimate the length of time that this immunity persists.

Prior to 1944 and the introduction of methods of rapid treatment, syphilis therapy extended over many months. This created the possibility, at least, of a higher degree of relative immunity among pre-1944 cases. To see what effect this would have on the incidence rates in this example, two extremes were considered. It was assumed that cases treated in the secondary stage were either all susceptible again or all immune. Cases treated in the primary stage have been indicated by Turner, Kluth, McLeod and Winsor [16] to be similar to non-syphilitic persons with regard to protective antibodies. The average difference between the two possible rates amounted to less than 2%. Therefore, it was decided to assume that all cases were susceptible after treatment.

attack rate would be given by this age frequency curve if the other two effects were constant.

No insight was available as to what type of trend to expect in yearly variations, so it seemed logical to fit seven parameters for year effects, one for each year. The relative effect of a calendar year on the observed syphilis cases may be interpreted as due to a variation in the real attack rate, administrative changes, or other uncontrolled elements. Study of the effects of calendar years becomes, therefore, a tool for the administrator in evaluating past programs of control.

It was postulated that the effect of yearly variation is multiplicative to the underlying age distribution. This is reasonable, since it was shown in a five-year study in New York State by Brumfield, Lade and Feldman [3] that an unfavorable year will result in a larger absolute increase among the modal age groups than it does among the other groups.

The effect of date of birth remains to be introduced in the model. Females born in the years 1912–1932 were eligible for this study. The date of birth may have an effect upon a person's chances of acquiring syphilis through education or other environmental factors. The present analysis is an attempt to measure the influence of such prior environment upon the probability of acquiring syphilis.

As was the case for year effects, no knowledge of a time trend was available, and, therefore, discrete parameters were fitted. Although there were 21 years, only seven parameters were used, one for each three-year period. This is the process of grouping previously mentioned in Section I to enable estimation of the terminal cohorts and to reduce the computational labor involved. A preliminary study of the data indicated that the loss of information by this technique would be extremely small.

Also, as was the case for year effects, it was assumed that the effects of the date of birth are multiplicative to the underlying age distribution. This means that if a particular three-year period was unfavorable with regard to syphilis, the contribution it would make to the increase of the attack rate would be proportional to the number already there because of age, and not an absolute amount disregarding age.

It was also assumed that the overall random error was proportional to the observed rate. This seemed justifiable in view of the multiplicative nature of the data.

With this background, it is now possible to formulate the mathematical model. If one lets y_{Xij} represent the observed rate occurring at age X, in the ith year, and therefore in the jth date of birth cohort,

difficulty is encountered because of zero values of y_{Xij}. The model allows y_{Xij} to be zero only at zero age and at infinite age. To overcome this limitation, the model was constructed relating $(y_{Xij}+1)$ to the effects of interest. The observed rate (plus unity) for any of the 105 cells in Table 1 may be expressed as a function of age, date of birth, year of discovery and random error as follows:

$$(3.1) \qquad y_{Xij} + 1 = X^{\beta_2} e^{\beta_0 X_0 + \beta_1 X + \sum_{i=1}^{7} \gamma_i X_i + \sum_{j=8}^{14} \delta_j X_j + \epsilon_{Xij}}$$

where

$X_0 = 1$ for all observations
X = age $(15 \leqq X \leqq 29)$
$X_1 = 1$ in 1941, 0 otherwise
$X_2 = 1$ in 1942, 0 otherwise
.
.
.
$X_7 = 1$ in 1947, 0 otherwise
$X_8 = 1$ for those born 1912–1914, 0 otherwise
$X_9 = 1$ for those born 1915–1917, 0 otherwise
.
.
.
$X_{14} = 1$ for those born 1930–1932, 0 otherwise

and β_0, β_1, β_2, are parameters in a Pearson Type III curve,
e = base of natural logarithms
γ_i = year effects
δ_j = date of birth or cohort effects
ϵ_{Xij} = random error, assumed normally and independently distributed with mean 0, and variance σ^2.

By taking logarithms, the right-hand side is expressible as a linear function of the parameters, and the mathematical model may now be rewritten

$$(3.2) \quad \log_e (y_{Xij}+1) = \beta_0 X_0 + \beta_1 X + \beta_2 \log_e X + \sum_{i=1}^{7} \gamma_i X_i + \sum_{j=8}^{14} \delta_j X_j + \epsilon_{Xij}.$$

There are 17 parameters β_0, β_1, β_2, $\gamma_1, \cdots, \gamma_7$, $\delta_8, \cdots, \delta_{14}$ which must be estimated. The method of least squares was used. The estimates of the parameters were designated by the corresponding Roman letters b, g, and d.

Before proceeding to the method of least squares, it was necessary to have assurance that the ϵ_{Xij} were normally and independently dis-

tributed. The requirement for a normal distribution is more readily satisfied than that of independence. The formation of a frequency distribution of the values of $\log_e (y_{xij}+1)$ showed that the assumption of a normal distribution is a plausible one.

In regard to the independence of the observations, however, one would expect the observations here to be serially correlated. This is true in most time-series data, especially when relating to economic or sociological material. That means that the observed value in any year will exercise an influence, up or down, upon the values of succeeding years.

The phenomenon of serial correlation is not unfamiliar to public health workers who long ago recognized the cycle in epidemics of measles and other diseases [4]. Although the nature of the correlation in syphilis is not as clear, there is little doubt that it is present. This is a serious violation of theory unless an attempt is made to estimate and remove the effects of this correlation.

If the effect of serial correlation is not removed, the F-test of the treatment mean square may be biased. (The term "treatment" is used here in a general sense and in this example would refer to years, date of birth or age.) The bias can be of such a nature that treatment effects are significant despite the fact that the expected value of the treatment mean square is less than the expected value of the error mean square.

It is currently believed that the fitting of parameters for year effects and date of birth effects removes most of the serial correlation [1]. It is realized that the present study may be criticized for the failure of the model to consider serial correlation over and above that removed by year and date of birth effects. Since, in ordinary least squares procedures, one must assume that deviations from the fitted surface are independent, an investigation of the residuals from the fitted surface was made. This showed non-significant correlations for lag 1 and 2.

IV. ANALYSIS OF VARIANCE IN OBSERVED INCIDENCE RATES

To make tests of significance upon the separate effects in the model, the sum of squares for each effect was adjusted for variations in the remaining variables. For example, the sum of squares associated with age was adjusted for variations both in years and date of birth. This type of adjustment is identical to that used in the analysis of variance with disproportionate sub-class numbers. The computational details have been omitted for the sake of brevity. It should be pointed out, however, that systematic short-cuts are available reducing the labor to a simplified routine [6], [7], [9].

The analysis of variance is given in Table 2. From the analysis of variance table, it is seen that age, calendar years and date of birth all displayed statistical significance in their influence upon the incidence of syphilis. The date of birth effect was significant at the 2 per cent level, whereas age and years were significant at the 1 per cent level. Age and calendar years accounted for approximately an equal amount of variation in the total sum of squares. Detailed descriptions of the characteristics of the effects of the three variables are contained in the following sections.

TABLE 2

ANALYSIS OF VARIANCE OF FACTORS INFLUENCING INCIDENCE RATE OF SYPHILIS FOR NEGRO FEMALES

Variance Source	d.f.	S.S.		M.S.	*F*-ratio	P
Total	105		591.3322			
Mean (Unadjusted)	1		508.7690			
Total Remaining	104		82.5632			
Years (Unadjusted)	6	9.0681	57.4112			
Age and Date of Birth (Adjusted)	8	48.3431				
Date of Birth (Unadjusted)	6	41.8734	57.4116			
Age and Years (Adjusted)	8	15.5382				
Age (Adjusted)	2	8.6309		4.3155	15.44	<0.01
Years (Adjusted)	6	8.2045		1.3674	4.89	<0.01
Date of Birth (Adjusted)	6	4.5383		0.7564	2.71	0.02
Remainder	90		25.1518*	0.2795		

$$* 82.5632 - \frac{57.4112 + 57.4116}{2} = 25.1518$$

V. INCIDENCE OF SYPHILIS ACCORDING TO AGE

In order to study the underlying age distribution of the incidence of syphilis, it is necessary to know the values of the three parameters in the Pearson Type III curve. The estimates of these parameters are b_0, b_1, b_2. Substituting them in (3.2), while holding years and date of birth constant, the result is

$$\log_e (y_{Xij} + 1) = -33.1611 - 1.0478X + 19.0208 \log_e X \quad (15 \leqq X \leqq 29). \tag{5.1}$$

Allowing X, the variable of age, to take all possible values in (5.1), the expected relative rates of incidence of syphilis according to age were calculated. The results are presented in Table 3 and Figure 2. It can be seen that the curve is similar to the one for non-white females reported by Vonderlehr and Usilton [26], but the central tendency occurs here at an age about five years earlier.

TABLE 3

ESTIMATED ANNUAL RATES OF INCIDENCE OF SYPHILIS FOR NEGRO FEMALES ACCORDING TO AGE WHEN ALL OTHER FACTORS ARE CONSTANT

Age	$\log_e (\hat{V}_{xij}+1)$	$\hat{V}_{xij}$ Cases per 1000 persons
15	2.631177	12.89
16	2.810980	15.63
17	2.916221	17.47
18	2.955650	18.21
19	2.936304	17.85
20	2.864081	16.53
21	2.744306	14.55
22	2.581354	12.21
23	2.379028	9.79
24	2.140753	7.51
25	1.869573	5.49
26	1.567768	3.80
27	1.237813	2.45
28	0.881610	1.41
29	0.501440	0.65

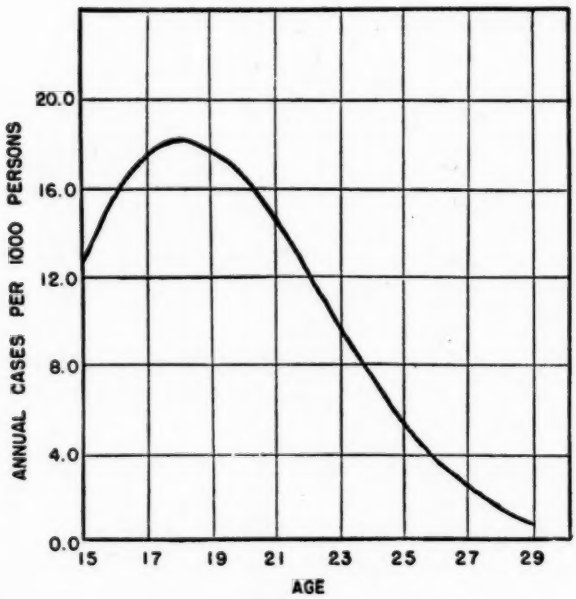


FIGURE 2

ESTIMATED ANNUAL AGE SPECIFIC RATES OF SYPHILIS INCIDENCE FOR NEGRO FEMALES IN STUDY GROUP, BY AGE

From Figure 2, it is clear that the age of maximum or modal incidence was the 17–19 years of age group. The graph shows how the distribution was skewed to the right and tapered off until a negligible incidence occurred at the 30-year mark. It serves as a reminder that venereal disease education should be started at the high school age if it is to be effective.

The exact age of maximum incidence was obtained from dividing 19.0208 by 1.0478, or 18.15 years. The 95% confidence interval[3] for this estimate is 15.4–20.7 years. This is one aspect of the age of maximum incidence. The subject will be treated in more detail in a later section.

It can be seen in Figure 2 that the underlying age distribution of the incidence rate of syphilis is relatively high at age 15. Contrariwise, the observed attack rates for this age were relatively low, except for 1946, as indicated in Table 1. This phenomenon is analogous to that pointed out by Frost for tuberculosis—viz., that age 20–30 was the most hazardous for tuberculosis in a male, although the observed mortality rates for persons aged 50–60 years in 1930 were higher than those for persons aged 20–30. The reason for this apparent but not real discrepancy is due to the effect of date of birth. In the present study, persons who were 15 years of age during 1941–1947 were members of the date of birth cohorts from 1926–1932. It will be shown in Section VII that this fact was beneficial to them and lowered the expected rates for persons of age 15 during 1941–1947.

VI. EFFECT OF YEARS UPON INCIDENCE RATE OF SYPHILIS

In accordance with the mathematical model expressed in (3.1), there were seven parameters for calendar years. The estimates of these parameters were designated by g_i $(i=1, \cdots, 7)$. The interpretation and units of these regression coefficients are governed by their use in the mathematical model. For example, g_k represents a multiplying factor, e^{g_k}, which was the result of the kth calendar year. The estimates of the regression coefficients were as follows:

[3] Computation of the 95% confidence interval involved solution for ξ_1 and ξ_2 in the following quadratic:

$$\frac{b_1 + \frac{b_2}{\xi}}{\sqrt{c_{11} + \frac{2c_{12}}{\xi} + \frac{c_{22}}{\xi^2}}} = t,$$

where t is the .05 Student t value at 90 degrees of freedom, and c_{ij} is an element in the variance-covariance matrix.

Year	Regression Coefficients g_i	e^{g_i} (Per Cent)
1941	−0.5429	58%
1942	−0.2843	75%
1943	−0.0590	94%
1944	0.2238	125%
1945	−0.0706	93%
1946	0.7271	207%
1947	0.0058	101%

From the tabulation of the results, it can be seen, for example, that the effect of the calendar year 1944 was to increase the observed rates 25 per cent over the average for the period 1941–1947. The average (100 per cent) would be the geometric mean of these seven yearly percentages. These results are presented graphically in Figure 3.

Inspection of the graph emphasizes the fact that 1946 was predominantly the heaviest year for the occurrence of syphilis. In fact, the effect of 1946 was to double the expected number of cases of syphilis. The explanations for this may be manifold. An essential one among them would be the return of service personnel from military duty. Such information should be invaluable for a health administrator.

It was predicted in 1945 that the calendar year 1946 would experience a large increase in the incidence rate of syphilis [13]. The present analysis is an attempt to quantify this increase while adjusting for the factors of age and date of birth.

It is interesting to investigate if the measurement of year effects without adjustment for age and date of birth gives similar results. If so, the additional computational labor would not be justifiable. Unadjusted percentage variations for each calendar year were computed by expressing the sum of the values of y_{xij} for each year as a percentage of the seven-year average of the sum of the values of y_{xij}. The unadjusted or crude effects of years upon the incidence of syphilis in the four observed counties are illustrated on Figure 3 for comparison.

Although similarity is evident, further examination reveals that the unadjusted effects are above and close to the adjusted effects during the first four years. For the last three years, however, the unadjusted percentage variations are below and farther away from the adjusted measurements. The reason for this is that with the increase of time, the date of birth effect is becoming more and more pronounced. With a larger series of data, this discrepancy would be more serious. Also, if

there were any sudden changes in the age composition of the group under study, further discrepancies would occur between the adjusted and unadjusted results.

VII. EFFECT OF DATE OF BIRTH UPON THE INCIDENCE OF SYPHILIS

Estimates of the date of birth effects in the mathematical model were designated by $d_8, \cdots, d_{14}$. The interpretation and units of these

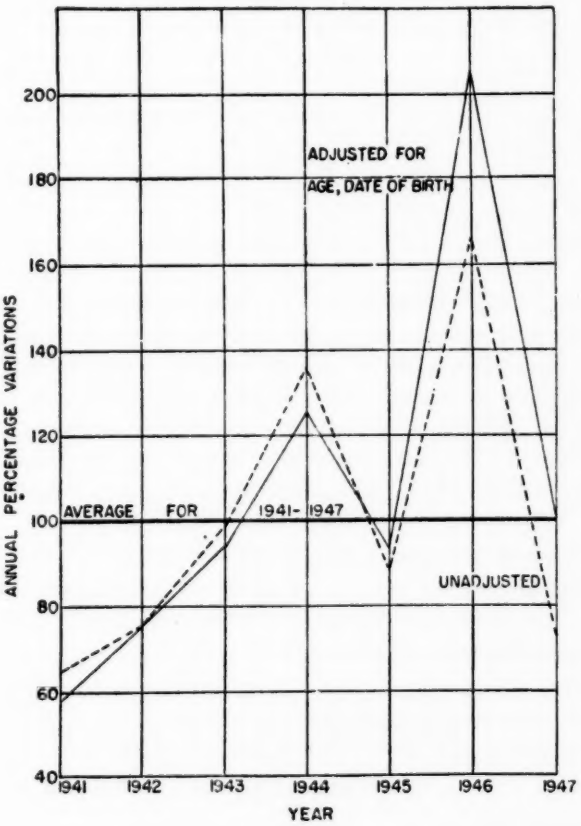


FIGURE 3

RELATIVE EFFECTS OF YEARS UPON SYPHILIS INCIDENCE FOR NEGRO FEMALES IN STUDY AREA, 1941–1947

regression coefficients are almost identical to those for calendar year effects. There is a slight exception in that the geometric mean of the percentages for the date of birth effects would not necessarily be

100%. This is due to the fact that there were unequal numbers of observations for each generation or date of birth period. For example, there were available six observations for those born in the period 1912–1914, fifteen observations for those born in the period 1915–1917, etc. A weighted geometric mean where the weights were 6, 15, 21, 21, 21, 15 and 6 respectively would result in a value of 100%. The estimates of the date of birth effects were calculated as follows, and are plotted on Figure 4.

Date of Birth	Regression Coefficients d_j	e^{d_j} (Per Cent)
1912–1914	0.8872	243%
1915–1917	0.1098	112%
1918–1920	−0.0966	91%
1921–1923	0.0287	103%
1924–1926	0.0506	105%
1927–1929	−0.1426	87%
1930–1932	−0.7446	47%

It is immediately apparent that there was a downward trend in the percentages with a temporary reversal for those born in the period 1921–1926, inclusive. That is, Negro females born more recently are experiencing less probability of acquiring syphilis. This may be explained by an increase in knowledge of prophylaxis during the period of time involved. The reason that this trend was temporarily reversed for those born from 1921–1926 should be a subject for further investigation.

The process of decrease appeared to follow that of a cubic rather closely. This hypothesis can be tested by means of a Student t, using functions of the estimated regression coefficients. Thus, to test linearity among the coefficients, one employs

$$t = \frac{(-3d_8 - 2d_9 - d_{10} + d_{12} + 2d_{13} + 3d_{14})}{\sqrt{\text{Estimated variance of numerator}}} \tag{7.1}$$

or

$$t = 0.99.$$

Hence, not significant.

In a similar manner, one may test the regression coefficients of date of birth effects for quadratic, cubic and quartic characteristics. This was done and the following t values obtained:

Quadratic:	t=0.34	Not significant
Cubic:	t=3.75	Highly significant
Quartic:	t=0.43	Not significant

It is evident that the tests of significance confirm the visual examination of the percentages themselves, namely that the decrease was according to a cubic curve. It is desirable to determine the coefficients of that cubic since the results can be used to refine the prediction of rates of syphilis incidence for the years of the study. It would substitute a continuous curve for the seven discrete parameters $d_8, \cdots, d_{14}$. It also would be of assistance in the next section where

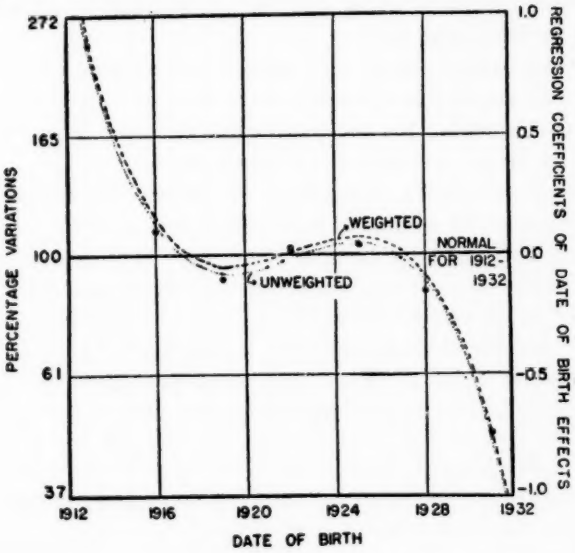


FIGURE 4

RELATIVE EFFECTS OF DATE OF BIRTH UPON SYPHILIS INCIDENCE FOR NEGRO FEMALES IN STUDY AREA, 1941-1947

an attempt is made to estimate the age of maximum incidence for each year during the study. Therefore, a cubic curve was fitted to the estimated regression coefficients as outlined in Appendix 2. The results of fitting two cubic curves, one weighted by variances and covariances and the other not, are indicated on Figure 4.

VIII. DETERMINATION OF ANNUAL AGE OF MAXIMUM INCIDENCE OF SYPHILIS

The underlying age distribution of the incidence of syphilis was depicted and evaluated in Section V. The seven-year overall age of maximum incidence was computed to be at 18.15 years. It was recognized, however, that this value fluctuated from year to year for various reasons.

Ruhland, Gillick and Chinn [13] reported that the age distribution of female admissions of primary and secondary syphilis to District of Columbia clinics went down during the war. They ascribed the result to changing social conditions and the tracing of contacts found through Selective Service testing. Similar investigations by Hirsh [11] and Rosenthal and Kerchner [12] also indicated a trend toward the teen ages by studying the median age of infection. Heller [10] employed a more refined technique in pointing out the inconsistency of this age shift both with locality and race.

The present analysis was an attempt to study changes in the modal age, or age of maximum incidence after year and date of birth effects had been considered. This was somewhat different from the problems discussed by these previous investigations.

First, by studying the logarithms of the estimated rates, the year effects of Section VI would be identical increments to each age group in the study and need not be considered here. The remaining two factors to be taken into account were the underlying age distribution and the date of birth effect. Interest was centered upon the manner in which they combined to produce the modal age each year.

The expected value of the incidence of syphilis when year effects were constant was expressed by combining (5.1) and (A.7)[3] so that

$$(8.1)\quad L_1 = \log_e(\vartheta_{Xij} + 1) = -33.1611 - 1.0478X + 19.0208\log_e X + 2.550157 - 0.55369907w + 0.05283556w^2 - 0.00157630w^3.$$

Let $v=1$ during 1941
$v=2$ during 1942
.
.
.
$v=7$ during 1947.

Substituting $w=(29+v-X)$ in (8.1), and taking the derivative with respect to age X, the maximum age was obtained from

$$(8.2)\quad \frac{dL_1}{dX} = -1.0478 + \frac{19.0208}{X} + (0.55369907) - 2(0.05283556)(29+v-X) + 3(0.00157630)(29+v-X)^2 = 0.$$

[3] This equation is derived in Appendix 2. It is a continuous curve expressing the date of birth effects.

As v took values 1, 2, $\cdots$, 7, the corresponding cubic equation in X was solved to determine the age of maximum incidence for each value of v. Each of the seven cubic equations obtained by this process was examined to ascertain that the solution of X yielded a true maximum.

The following maxima were calculated for each year:

Year	Age of Maximum Incidence
1941	17.65
1942	17.85
1943	18.12
1944	18.46
1945	18.84
1946	19.28
1947	19.76

It can be noted that the age of maximum syphilis incidence increased during and after the war years. In the quoted studies [11], [12], [13], it was pointed out that the age distribution went down. Those investigations do not conflict with the results of this paper since the peak age was studied here. In the former instances, the age distribution went down as a result of relatively larger absolute increases among the modal groups (i.e., young persons) during the war years. This phenomenon coincided with the multiplicative nature of the model used here.

SUMMARY

An attempt was made to present a method of analyzing simultaneously some of the important factors affecting the incidence of syphilis in a group specific for race and sex. A mathematical model was constructed which took into account three important, measurable factors—age, date of birth and year under observation. The technique was illustrated by using data on syphilis incidence among Negro females in four counties during a seven-year period. Several instances were indicated where erroneous inferences were possible if the observed rates of syphilis incidence were analyzed separately for each factor instead of considering their effects simultaneously.

For the Negro females, all three factors were significant in a statistical sense. The underlying age distribution of syphilis incidence according to age was determined. It was observed that the Negro females experienced a peak at an early age, i.e., 18.15 years.

Year effects were measured as percentage changes from the average for the period. The results showed that the year following the return of service personnel (1946) was the year of maximum incidence, and had, in fact, doubled the average rate of incidence for the period 1941–1947.

The date of birth effects were also estimated. It was found that for those born in the period 1912–1932, the ones born more recently have a smaller probability of acquiring syphilis. There was a temporary reversal of this trend for the group born from 1921–1926.

Finally, the age of maximum incidence of syphilis was computed for each year. It was found that the age of maximum incidence for these females increased slightly during the war years due to the effect of date of birth.

APPENDIX 1

Population figures for postcensal years on a geographical, sex, race and age basis were necessary to determine the respective exposed populations in this problem. It was realized that an exact knowledge of the exposed populations would be unobtainable because of the sparse information available, especially for the war years. Procedures for estimating the current population of counties have been suggested in [22] and summarized recently by Shryock and Lawrence [15]. These methods were not utilized, however, because of the unreliability of some of the local information, especially school populations. The most feasible technique was used, therefore, to determine the county population while recognizing that the results were crude estimates. It is believed, nevertheless, that the final analysis would not be materially affected by any allowable discrepancy in the population.

Postcensal estimates for counties and cities, which are sporadically furnished by the Bureau of the Census, were used as a basis. These estimates by this federal agency were based upon sample surveys, school enrollments, knowledge of vital statistics in conjunction with immigration, and registration for war ration books.

Each county was treated here as a separate entity. The county results were not combined until the last step, since it was apparent that county variations in all factors would make such an estimate more accurate than a combination at the outset. The work was divided essentially into four parts, each of which is herewith discussed in detail.

Part 1

The first step was to estimate the total population in each of the four counties, as of July 1, for the years 1941–1947, inclusive. To

facilitate the estimation of county population, by years, the following information was available from publications of the Bureau of the Census:

1940: A complete census of the population within each of the four counties by sex, race and age, as of April 1, is available. The age grouping is in units of five years, or more, for those over 20 years of age. For 15–19 years, it is possible to obtain separate totals for each of the five years. The county total is also classified to indicate the type of population, i.e., rural farm, rural non-farm, and urban. [23]

1942: An estimate of the total population as of May 1, by counties, is available. [20]

1943: An estimate of the total population, by counties, as of March 1 [19] and November 1 [21], is available.

1947: An estimate of the changes in population of the United States from 1940 to 1947, for white and non-white, by region and type of population, is available. [17]

1940–1947: An estimate of the population in North Carolina for each year as of July 1, is available. [18]

Data was also made available by Selective Service authorities listing the number of males who entered the armed services by year, county, race and age. This permitted a correction of the wartime postcensal estimates which included civilian residents only and excluded all persons in military service.

The change in the four counties from 1940 to 1942 paralleled somewhat the increase made by the state during that same period. It was assumed, therefore, that the total population in each of the four counties for 1941 was in the same ratio to 1940 and 1942 as was that of the state as a whole.

To estimate the county populations for 1944–1947 inclusive, it was assumed that the change in population from 1940 to 1947 in these four counties, by race and type of population, was similar to the change in such groups in the continental United States. By using the 1940 Census data and its classification of county populations by race and type within each county, it was possible to estimate the number of persons in the same race and type groups for each county in 1947.

It was observed that the change in total population of the counties from 1943–1947 was not dissimilar to the change in the entire state

for that same period. Consequently, the county populations were graduated, percentage-wise, from 1943–1947, in conformity with the changes for the entire state. This completed an estimate for each county for the years 1941–1947, inclusive.

Part 2

The county populations were classified by race and sex. There were two years, 1940 and 1947, in which an estimate of the percentage of white and Negro persons in each county was obtainable. These percentages, at the two terminal years, differed only slightly, and it seemed reasonable to graduate those differences throughout the period.

The sex ratio in each county among the Negro population was assumed to be the same as in the 1940 Census. This seemed justifiable in view of the fact that the population estimates had been adjusted for the males in military service.

Part 3

The numbers of non-white females in each county were classified according to five-year age groups. The procedure was as follows:

The 1940 Census indicates the proportion of the county population in each age group by a sex and race classification. This proportion varies annually as a result of fluctuating birth and mortality rates. The Bureau of the Census estimated the 1947 composition of the population by age groups, for each sex and race, and by region. By comparing these figures with analogous ones for the South in 1940, an adjustment factor was derived.

For example, consider county A which had 11.9 per cent of the non-white females in age group 15–19 in 1940. The percentage of non-white females in that same age group for the entire South was 10.8 per cent in 1940 and 9.8 per cent in 1947. Therefore, the estimate of the percentage of non-white females in that age group in county A in 1947 is given by $(9.8/10.8)\ (11.9) = 10.8$.

By this process, one knew the number of non-white females in each county in age groups 15–19, 20–24, and 25–29 in 1940, and had an estimate of the same groups in 1947. The differences between the two terminal years were graduated for the intervening years. Thus, the number of Negro females was estimated for age groups 15–19, 20–24, 25–29, etc., in each county in each year of the study.

Part 4

The next step was the classification of the five-year age groups

into one-year increments. The procedure adopted was to obtain estimated populations at the pivotal values 17, 22, and 27 by using King's Central Difference Formula [27]. Following this, an osculatory interpolation process in the form of the Karup-King formula was used. These formulas are summarized in pages 123–126 of a publication by the Bureau of the Census [24].

APPENDIX 2

Fitting a Cubic Curve to Estimated Date of Birth Regression Coefficients

To determine the cubic curve which appears to underlie the estimated regression coefficients, the problem of considering the weights to be assigned each of the seven estimates arose. It is well known that the proper system of weights is one which is determined by the invariance, or reciprocals of the variance. In this case there were two complications. The first was that there were covariances as well as variances. The second was that d_{14} had zero variance and covariance with all other estimated regression coefficients. This implied that the curve must be anchored at d_{14}, since its weight was infinite.

First, one may consider the general problem.[4] Assume that there are p regression coefficients $\beta_1, \cdots, \beta_p$ whose estimates are given by $b_1, \cdots, b_p$. The sample variance-covariance matrix will be c_{ij} and its inverse will be c^{ij} $(i, j=1, 2, \cdots, p)$. In the case where the X's are fixed, the estimates of the regression coefficients are multivariate normally distributed and one may consider

$$L = \text{Log (Joint Distribution of } b\text{'s)} = -\tfrac{1}{2} \sum_{i,j=1}^{p} c^{ij}(b_i - \beta_i)(b_j - \beta_j). \tag{A.1}$$

The curve to be estimated is

$$\beta_i = \alpha_0 l_{0i} + \alpha_1 l_{1i} + \alpha_2 l_{2i} + \alpha_3 l_{3i} \tag{A.2}$$

where α_0, α_1, α_2, α_3 are regression coefficients in the cubic equation, and

$l_{0i} = 1$ for all observations

l_{1i} = an equally-spaced variable located at the midpoint of each three-year date of birth period

$l_{2i} = l^2_{1i}$

$l_{3i} = l^3_{1i}$.

Designate by a_0, a_1, a_2, a_3 the estimates of the regression coefficients

[4] This method of approach was suggested to the authors by Professor W. G. Cochran, School of Hygiene and Public Health, Johns Hopkins University.

in (A.2). In order to evaluate these estimates by maximum likelihood, one may write

$$(A.3) \qquad L = -\tfrac{1}{2}\sum_{i,j}^{p} c^{ij}(b_i - \alpha_0 t_{0i} - \alpha_1 t_{1i} - \alpha_2 t_{2i} - \alpha_3 t_{3i})(b_j - \alpha_0 t_{0j} - \alpha_1 t_{1j} - \alpha_2 t_{2j} - \alpha_3 t_{3j}).$$

By setting $\partial L/\partial \alpha_u = 0$ $(u = 0, 1, 2, 3)$, one obtains four sets of normal equations for the solution of estimates of the α's.

For example, if $\partial L/\partial \alpha_0 = 0$, it clearly results in the first equation. The only unknowns in this equation are the α's. That is, one obtains

$$(A.4) \qquad \sum_{i,j}^{p} c^{ij}(b_i t_{0j} + b_j t_{0i}) = \sum_{i,j}^{p} c^{ij}\{2\alpha_0 t_{0i} t_{0j} + \alpha_1(t_{1i} t_{0j} + t_{1j} t_{0i}) + \alpha_2(t_{2i} t_{0j} + t_{2j} t_{0i}) + \alpha_3(t_{3i} t_{0j} + t_{3j} t_{0i})\}.$$

Examination of (A.4) revealed that the amount of computational labor would be prodigious. Besides inverting a matrix of order 6, the summation and multiplication for each normal equation made the problem unwieldy. There were two saving features in the process. The first was that t_{1i} was equally spaced and systematic short cuts could be introduced. The second was that since the problem required the curve to be anchored at d_{14}, the need for α_0 was eliminated by placing that point at the origin (0, 0).

By assuming d_{14} was at the origin, the corresponding value for t_1 was zero. It also involved translating the values of d_j to the right 0.7446 units. To pass the curve through the origin, the uncorrected sums of squares and products were used in the three equations to estimate α_1, α_2, α_3.

Following through these computations, the equation to predict the date of birth effect, D, was given by

$$(A.5) \qquad D = -0.995510t - 0.375682t^2 - 0.042560t^3.$$

By making the substitution $t = (w - 20)/3$ in (A.5), the curve was shifted so that D represented the effect of birth in year w upon the incidence of syphilis. That is,

$w =$ 1 for 1912

$w =$ 2 for 1913

$w =$ 3 for 1914

(A.6) . .
. .
. .

$w = 21$ for 1932

Carrying through this substitution resulted in

(A.7) $D = 2.550157 - 0.55369907w + 0.05283556w^2 - 0.00157630w^3$.

To compare the fitted curve with the seven estimated regression coefficients, D was determined for values of $w=2, 5, 8, \cdots, 20$. Thus, the following comparison was made to observe the closeness of fit by the cubic equation (A.7).

w	D	Translated Regression Coefficients
2	1.6415	1.6318
5	0.9055	0.8544
8	0.6950	0.6480
11	0.7545	0.7733
14	0.8288	0.7952
17	0.6624	0.6020
20	0	0

The cubic equation (A.7) is plotted on Figure 4.

One may ponder if the extra labor in fitting equation (A.2) by a weighted scheme resulted in a cubic much different from one that would have been obtained by an equal-weighted regression. The latter was calculated readily with the aid of orthogonal polynomials and the result was

(A.8) $D_1 = 1.772771 - 0.555156w + 0.052805w^2 - 0.00157064w^3$.

This equation is also included in Figure 4. It will be noticed that there was a strong similarity between (A.7) and (A.8). Nevertheless, differences did exist, especially in the coefficients of the lower powers of w. Examination of the graph in Figure 4 reveals that the weighted curve, or proper cubic equation, was closer to the estimated regression coefficients only at both extremes and at the median value.

BIBLIOGRAPHY

[1] Anderson, R. L. (1949). Personal Communication.
[2] Birnbaum, Z. W., and Zuckerman, Herbert S. (1948). On the Determination

of the Dependence of a Disease, Especially Cancer, on Age. *University of Washington Publications in Mathematics*, Vol. 3, No. 1, July.

[3] Brumfield, William A., Jr.; Lade, James H.; and Feldman, Louis (1942). The Epidemiology of Syphilis Based Upon Five Years' Experience in an Intensive Program in New York State. *American Journal of Public Health*, Vol. 32, No. 8, August, 793–802.

[4] Chassan, Jacob B. (1948). The Autocorrelation Approach to the Analysis of the Incidence of Communicable Diseases. *Human Biology*, Vol. 20, No. 2, May, 90–108.

[5] Dunn, Halbert L. (1947). Health and Social Statistics for the City. *American Journal of Public Health*, Vol. 37, No. 6, June, 739–743.

[6] Dwyer, Paul S. (1941). The Solution of Simultaneous Equations. *Psychometrika*, Vol. 6, No. 2, 101–129.

[7] Fisher, R. A. (1948). *Statistical Methods for Research Workers. Tenth Edition.* Oliver and Boyd, Edinburgh, Section 29.

[8] Frost, Wade Hampton (1939). The Age Selection of Mortality from Tuberculosis in Successive Decades. *American Journal of Hygiene*, Vol. 30, No. 3, Section A, 1–96.

[9] Greenberg, B. G. (1949). Unpublished doctoral dissertation, North Carolina State College.

[10] Heller, J. R., Jr. (1946). Wartime Changes in the Age Distribution of Females Infected with Syphilis. *American Journal of Public Health*, Vol. 36, No. 5, May, 501–509.

[11] Hirsh, Joseph (1943). Boomtown Wins a Battle. *Journal of Social Hygiene*, Vol. 29, No. 9, 583–593.

[12] Rosenthal, Theodore, and Kerchner, George (1944). Trend in Age of Acquiring Venereal Disease in New York City, 1940–1943. *Journal of Venereal Disease Information*, Vol. 25, No. 12, 361–365.

[13] Ruhland, George C., Gillick, Frederick C., and Chinn, Ben D. (1945). The Possibility of Predicting the Future Needs in Venereal Disease Control. *Journal of Venereal Disease Information*, Vol. 26, No. 10, October, 222–229.

[14] Shryock, Henry S., Jr. (1947). Postcensal Population Data for Cities. *American Journal of Public Health*, Vol. 37, No. 11, November, 1455–1460.

[15] Shryock, Henry S., Jr., and Lawrence, Norman (1949). The Current Status of State and Local Population Estimates in the Census Bureau. *Journal of American Statistical Association*, Vol. 44, No. 246, June, 157–173.

[16] Turner, Thomas B., Kluth, Fred C., McLeod, Charlotte, and Winsor, Charles P. (1948). Protective Antibodies in the Serum of Syphilitic Patients. *The American Journal of Hygiene*, Vol. 48, No. 2, September, 173–181.

[17] United States Bureau of the Census, *Current Population Reports*, Population Characteristics, Series P-20, No. 9, January 19, 1948.

[18] United States Bureau of the Census, *Current Population Reports*, Population Estimates, Series P-25, No. 12, August 9, 1948, Table 3.

[19] United States Bureau of the Census, *Population*, Series P-3, No. 38, October 31, 1943, Table 3.

[20] United States Bureau of the Census, *Population*, Series P-3, No. 33, February 25, 1943, Table 3.

[21] United States Bureau of the Census, *Population Special Reports*, Series P-44, No. 3, February 15, 1944, Table 3.

[22] United States Bureau of the Census, *Population Special Reports*, Series P-47, No. 4, April 30, 1947.

[23] United States Bureau of the Census, *Sixteenth Census of the United States.* Characteristics of the Population, North Carolina, Second Series, Table 22.

[24] United States Bureau of the Census, *United States Life Tables and Actuarial Tables, 1939–1941.* Washington, D. C., 1946, 119.

[25] Vonderlehr, R. A., and Usilton, Lida J. (1938). The Chance of Acquiring Syphilis and the Frequency of Its Disastrous Outcome. *Journal of Venereal Disease Information*, 19, November, 1–9.

[26] Vonderlehr, R. A., and Usilton, Lida J. (1943). The Extent of the Syphilis Problem at the Beginning of World War II. *New York State Journal of Medicine*, Vol 43, No. 19, October 1, 1825–1829.

[27] Whittaker, E. T., and Robinson, G. (1929). *Calculus of Observations, Second Edition*, Blackie and Son, Ltd., London, 57–61.

[28] Wright, John J. (1948). Field Epidemiological Study of Syphilis. *The Health Bulletin*, North Carolina State Board of Health, Vol. 63, No. 6, June, 15–16.

[29] Wright, John J., and Sheps, Cecil G. (1949). Reports of the North Carolina Syphilis Studies. I. An Evaluation of Case-Finding Measures in Syphilis Control. *Journal of Venereal Disease Information*, Vol. 30, No. 2, February, 35–52.

LABOR-FORCE PARTICIPATION AND EMPLOYMENT OPPORTUNITIES FOR WOMEN

NEDRA BARTLETT BELLOC
Institute of Industrial Relations, University of California, Berkeley

Studies of labor-force participation have heretofore stressed the influence of earnings as a determinant of the proportion of the population that seeks employment. Low earnings are associated with high propensities to engage in the labor force. The variation in labor-force participation is greater among females than among males in the large cities of the United States. This study points out that high female labor-force participation rates are found in communities in which large numbers of jobs suitable for women are found in textile, apparel, tobacco, and electrical machinery manufacturing. Conversely, low labor-force participation among women is found in communities in which the dominant industries are iron and steel, machinery, automobile, and basic rubber manufacturing. It is the thesis of this paper that the availability of employment opportunities for women is a major factor in determining labor-force participation.

EFFECTIVE national planning for full employment requires knowledge of the factors affecting the proportion of the population that works or seeks work. If, as suggested by Clarence D. Long,[1] the labor force tends to be a fairly stable proportion of the population, forecasts can be made with some confidence regardless of fluctuations in the business cycle or movements of real earnings. If, on the other hand, participation in the labor force is sensitive to economic changes, much work remains to be done in measuring the effects of these and other factors influencing the size and composition of the labor force. It is the thesis of this study that the factor of employment opportunities for women is a major determinant of the variation which exists in the proportion of the population in the labor force.

It has been pointed out by Paul H. Douglas[2] and more recently by Long[3] that the proportion of the population in the labor force is smaller in areas where real earnings are higher. In addition to earnings, this paper examines two other factors which correlate with labor-force participation, namely: employment opportunities for women, and the

[1] "The Labor Force and Economic Change," a chapter in *Insights into Labor Issues*, R. A. Lester and J. Shister, eds. (New York: Macmillan Co., 1948), p. 353.

[2] *The Theory of Wages* (New York: Macmillan Co., 1934), Chap. XI; Erika H. Schoenberg and Paul H. Douglas, "Studies in the Supply Curve of Labor," *The Journal of Political Economy*, XLV (February, 1937), pp. 45–79.

[3] *Op. cit.*, p. 338.

percentage of the population white. The data used are those given in the 1940 U. S. Census of Population for the 91 cities over 100,000 population, excluding only Washington, D. C. (See Table 1.)

VARIATION IN LABOR-FORCE PARTICIPATION

The proportion of the population over 14 years of age in the labor force varied from 47.8 per cent in Long Beach, Calif., to 62.5 per cent in Charlotte, N. C. (Mean, 55.5; S. D., 2.5). However, when the data are separated by sex, it is apparent that participation of females varied much more than that of males. Male participation varied from 75.3 per cent in Long Beach to 84.7 per cent in Detroit (Mean, 80.1; S. D., 2.1) and female participation varied from 20.2 per cent in Gary, Indiana, to 44.1 per cent in Charlotte (Mean, 32.4; S. D., 4.4).

Since participation varies with age, the influence of differing age distributions in the various cities was eliminated by adjusting the labor-force participation rates, using the population of Chicago as the standard. This adjustment reduced the variation in the participation rates of males (S. D., 1.6 per cent), but made no change in the variation for females.

RELATIONSHIP OF EARNINGS TO LABOR-FORCE PARTICIPATION

The coefficient of correlation between male earnings[4] and the participation of females in the labor force was −.54, confirming the findings of research with other data. Douglas explained this phenomenon by saying, "Where wages are low (women) will be driven into the labor market in much greater numbers in order to eke out the family income than where the earnings of their husbands are higher."[5] Long's comments on this point follow: "Propensities . . . have risen in the case of women, the rise having been especially great among married women. The variations of the individual groups do not seem closely associated in time with income or indeed with any other single factor. . . . Women entered the labor force in spite of the large rises in real income. Indeed these income rises should, on the basis of the static

[4] The median earnings of males who worked 12 months in 1939. This measure gave higher correlations than earnings of all males. Earnings of females showed low correlations with all factors studied except percentage of population white. (See Table 3.)

Real earnings for 29 cities were calculated as follows: the total cost of goods and services for a family of four in these cities in March 1946 ("The City Worker's Family Budget," *Monthly Labor Review*, 66 [February, 1948], 133–170) was projected back to August 1939 by using the percentage change in the Consumers' Price Index from August 1939 to March 1946 (*Monthly Labor Review*, 62 [May, 1946], 792). The estimated variation in the cost of living thus obtained enabled conversion of 1939 money earnings to real earnings. The coefficient of correlation between money earnings and real earnings in the 29 cities was .98.

[5] *The Theory of Wages*, p. 276.

TABLE 1

WAGE OR SALARY INCOME IN 1939 AND LABOR FORCE PARTICIPATION IN 1940 IN 91 CITIES OVER 100,000 POPULATION

Key*	City	Median Wage or Salary Income in 1939 of Experienced Persons in the Labor Force in 1940†				Percentage of Population Over 14 Years of Age in Labor Force in 1940‡	
		Total	Males	Females	Males Who Worked 12 Months	Males	Females
1	Birmingham, Ala.	$ 623	$ 686	$ 272	$1185	82.7	31.2
2	Long Beach, Calif.	880	1089	438	1438	78.0	25.5
3	Los Angeles	844	981	629	1339	78.3	32.9
4	Oakland	1062	1263	699	1615	79.8	29.4
5	Sacramento	1004	1175	810	1588	80.2	34.7
6	San Diego	730	842	417	1205	78.8	27.2
7	San Francisco	1062	1212	862	1524	78.8	35.9
8	Denver, Colo.	788	962	559	1318	79.2	31.3
9	Bridgeport, Conn.	862	1049	644	1319	83.7	37.5
10	Hartford	1002	1170	769	1397	82.0	37.9
11	New Haven	805	990	612	1272	82.2	37.1
12	Wilmington, Del.	892	1080	571	1403	83.1	35.0
13	Jacksonville, Fla.	582	715	334	1013	81.1	36.0
14	Miami	655	814	412	1186	79.6	38.1
15	Tampa	484	589	331	860	81.8	40.8
16	Atlanta, Ga.	577	718	369	1042	82.7	41.1
17	Chicago, Ill.	893	1063	656	1391	82.0	33.5
18	Peoria	972	1238	576	1482	83.0	31.0
19	Fort Wayne, Ind.	944	1203	613	1515	81.6	29.7
20	Gary	1201	1316	471	1555	82.6	18.8
21	Indianapolis	873	1040	613	1338	82.3	32.1
22	South Bend	914	1197	562	1517	83.8	30.7
23	Des Moines, Iowa	906	1093	668	1428	81.3	32.7
24	Kansas City, Kan.	764	939	467	1282	82.9	28.8
25	Wichita	694	861	403	1200	80.4	30.2
26	Louisville, Ky.	706	850	477	1148	83.0	33.2
27	New Orleans, La.	563	686	335	950	82.0	33.6
28	Baltimore, Md.	802	1003	525	1239	81.4	33.6
29	Boston, Mass.	857	992	700	1320	79.6	35.5
30	Cambridge	909	1080	708	1379	80.0	37.9
31	Fall River	636	725	527	1032	82.6	45.2
32	Lowell	701	797	518	1172	81.0	37.3
33	New Bedford	626	721	490	1056	81.9	42.2
34	Somerville	969	1136	722	1439	82.7	31.1
35	Springfield	937	1161	649	1431	82.2	33.6
36	Worcester	954	1169	628	1488	78.9	31.4
37	Detroit, Mich.	1034	1207	613	1634	84.1	27.5
38	Flint	1233	1443	568	1747	83.6	26.3
39	Grand Rapids	853	1029	561	1323	81.9	29.1
40	Duluth, Minn.	848	1033	543	1450	81.0	28.8
41	Minneapolis	908	1164	661	1492	80.7	34.7
42	St. Paul	923	1157	653	1471	81.5	32.5
43	Kansas City, Mo.	747	922	514	1298	83.6	35.7
44	St. Louis	747	916	527	1264	83.6	34.5
45	Omaha, Neb.	808	979	538	1306	82.0	31.5

TABLE 1—*Continued*

Key*	City	Median Wage or Salary Income in 1939 of Experienced Persons in the Labor Force in 1940†				Percentage of Population Over 14 Years of Age in Labor Force in 1940‡	
		Total	Males	Females	Males Who Worked 12 Months	Males	Females
46	Camden, N. J.	$ 786	$ 965	$ 561	$1312	84.1	33.8
47	Elizabeth	1012	1212	635	1481	83.3	32.7
48	Jersey City	975	1134	720	1430	84.1	33.8
49	Newark	815	993	622	1329	81.9	34.1
50	Paterson	673	777	507	1247	83.8	35.8
51	Trenton	748	916	544	1266	78.2	36.8
52	Albany, N. Y.	910	1072	697	1379	81.8	36.0
53	Buffalo	826	1022	540	1421	81.2	29.0
54	New York City	885	1040	686	1390	81.0	33.8
55	Rochester	952	1148	708	1484	79.7	35.5
56	Syracuse	857	1039	610	1371	80.5	31.9
57	Utica	703	861	506	1237	79.9	35.2
58	Yonkers	1036	1212	732	1577	81.8	32.8
59	Charlotte, N. C.	618	756	368	1057	83.4	41.3
60	Akron, Ohio	966	1168	634	1578	81.8	26.8
61	Canton	925	1117	517	1419	83.2	25.5
62	Cincinnati	852	1044	604	1348	81.2	31.4
63	Cleveland	825	996	580	1398	81.8	30.4
64	Columbus	874	1042	618	1328	76.7	31.3
65	Dayton	1016	1221	629	1498	81.9	30.1
66	Toledo	828	1039	538	1462	81.0	30.0
67	Youngstown	926	1105	476	1492	82.7	24.2
68	Oklahoma City, Okla.	672	832	400	1239	79.7	32.7
69	Tulsa	695	912	402	1345	81.7	32.4
70	Portland, Ore.	860	1061	577	1487	80.6	32.9
71	Erie, Pa.	819	1020	482	1368	82.2	27.4
72	Philadelphia	772	950	569	1303	82.1	34.4
73	Pittsburgh	819	996	561	1400	81.5	28.4
74	Reading	691	841	515	1207	84.0	39.2
75	Scranton	631	719	491	1322	81.3	29.2
76	Providence, R. I.	710	821	579	1203	83.4	38.7
77	Chattanooga, Tenn.	653	755	409	978	82.3	34.7
78	Knoxville	622	712	447	968	82.2	36.2
79	Memphis	562	704	294	987	83.5	36.1
80	Nashville	564	687	361	983	81.9	37.2
81	Dallas, Tex.	638	779	403	1127	82.9	38.4
82	Fort Worth	593	751	341	1144	82.2	33.2
83	Houston	734	933	387	1255	82.3	33.0
84	San Antonio	425	503	309	705	82.0	30.7
85	Salt Lake City, Utah	940	1148	622	1470	79.8	26.1
86	Norfolk, Va.	663	812	348	1162	84.0	31.6
87	Richmond	710	890	497	1204	81.8	42.3
88	Seattle, Wash.	918	1125	624	1555	79.5	32.3
89	Spokane	790	973	509	1398	81.1	30.2
90	Tacoma	913	1092	526	1450	81.4	26.0
91	Milwaukee, Wis.	934	1157	639	1498	81.4	30.4

* See Figure 1.

† Calculated from Sixteenth Census of the United States, 1940, *Population*, Vol. III, Table 15 for each state.

‡ Calculated from Table 5 for each state. Standardized to the age distribution of the Chicago population in 1940.

association, have suggested their large-scale exit. The paradoxical rise has been due no doubt to the great social upheaval that occurred over the decades and made it at once feasible, respectable, and possibly compulsory, for women to work gainfully."[8]

RACIAL COMPOSITION OF THE POPULATION

The percentage of the population white varied from 58.5 in Memphis, Tenn., to 99.9 in Lowell, Mass. In general, higher labor-force participation rates were found in the South than in the North and West. However, some of the New England cities were outstanding exceptions, showing high labor-force participation in predominately white communities. The percentage of the population white showed a coefficient of −.26 with participation of women in the labor force, and .57 with male earnings.

EMPLOYMENT OPPORTUNITIES FOR WOMEN

Unfortunately, there appears to be no really adequate measure of employment opportunities for women. The "core" of women's em-

TABLE 2

COEFFICIENTS OF CORRELATION BETWEEN FEMALE LABOR-FORCE PARTICIPATION, MALE EARNINGS, PERCENTAGE OF FEMALES EMPLOYED IN MANUFACTURING AND DOMESTIC SERVICE, AND PERCENTAGE OF POPULATION WHITE IN 91 U. S. CITIES OVER 100,000 POPULATION, 1940

Factor	Coefficient of correlation with factor number			
	(1)	(2)	(3)	(4)
(1) Percentage of females over 14 years of age in the labor force (standardized for age)	—	−.54	.54	−.26
(2) Median earnings of males who worked 12 months in 1939	−.54	—	−.42	.57
(3) Percentage of females employed in manufacturing and domestic service	.54	−.42	—	−.29
(4) Percentage of population white	−.26	.57	−.29	—

Levels of significance:

P	.05	.02	.01
r	.206	.244	.269

ployment lies in particular fields: clerical work, teaching, nursing, selling. Occupations in these fields, together with some domestic service and manufacturing, provide employment for a certain proportion of the girls and women in the population. If women are employed in considerably larger numbers, they are likely to be found either in

[8] *Op. cit.*, p. 352.

manufacturing or domestic service. The percentage of women employed in these two fields varied from 14.1 in Sacramento, California to 74.1 in Fall River, Massachusetts. In the absence of a better measure, this percentage was used to indicate variation in employment opportunities for women. The coefficients of correlation between this measure and the other factors are summarized in Table 2. Other coefficients are shown in Table 3.

TABLE 3

COEFFICIENTS OF CORRELATION BETWEEN EARNINGS AND OTHER FACTORS RELATED TO LABOR FORCE PARTICIPATION IN 91 CITIES OVER 100,000 POPULATION

Factor	Coefficients of Correlation with			
	Median Wage or Salary Income in 1939 of Experienced Persons in Labor Force in 1940*			Percentage of Population White
	Males Who Worked 12 Months	All Males	All Females	
Percentage of male population over 14 years of age in labor force in 1940—standardized for age*‡	−.14	−.11	−.02	−.26
Percentage of female population over 14 years of age in labor force in 1940—standardized for age*§	−.54	−.50	−.05	−.26
Percentage of male population over 14 years of age in labor force in 1940†	−.12	−.07	−.24	−.52
Percentage of female population over 14 years of age in labor force in 1940†	−.57	−.53	−.14	−.39
Percentage of population white	.57	.53	.65	—
Percentage of females employed in manufacturing and domestic service	−.42	−.35	−.22	−.29

Levels of significance:

P	.05	.02	.01
r	.206	.244	.269

* Measures calculated are shown in Table 1. Other measures were taken directly from the Sixteenth Census of the United States, 1940.

† Coefficient between percentage of total population over 14 years of age in labor force and median earnings of all workers: −.36.

‡ Coefficient with percentage of males employed in manufacturing: .43.

§ Coefficient with percentage of females employed in manufacturing: .38.

The percentage of females employed in manufacturing and domestic service correlated highly (.54) with the percentage of females over 14 years of age in the labor force. Lesser, but significant, negative correlations were shown with male earnings and percentage of population white, the coefficients being −.42 and −.29, respectively.

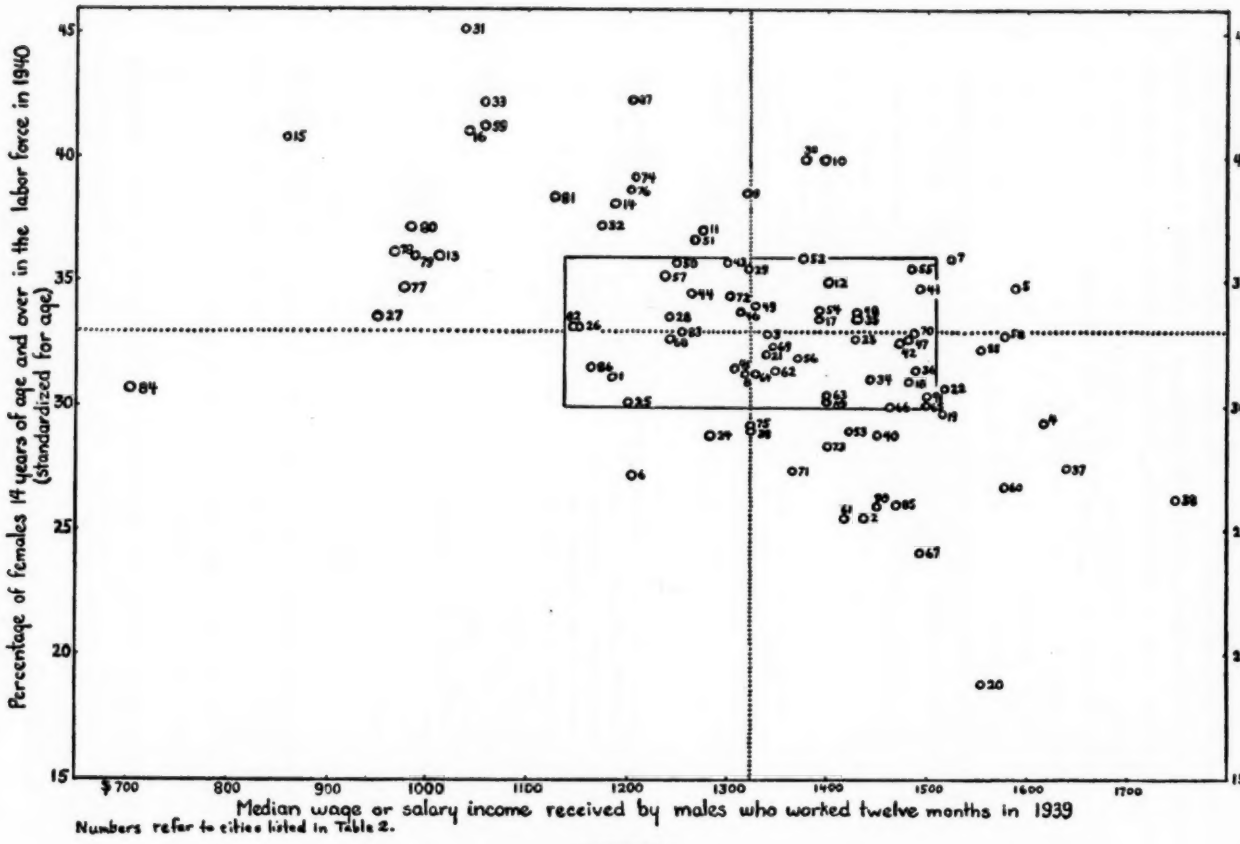


FIGURE 1

RELATIONSHIP BETWEEN MALE EARNINGS AND FEMALE PARTICIPATION IN THE LABOR FORCE, NINETY-ONE LARGE CITIES IN THE UNITED STATES: 1940.

The coefficients of partial and multiple correlation (using the notation of Table 1), were as follows:

$r_{12.3}$	−.42	$r_{12.4}$	−.50	$r_{23.4}$	−.32	$r_{13.4}$	.50
$r_{13.2}$	.41	$r_{14.2}$	.07	$r_{34.2}$	−.08	$r_{14.3}$	−.13
$r_{23.1}$	−.17	$r_{24.1}$	.53	$r_{24.3}$	.52	$r_{34.1}$	−.19
				$R_{1.23}$	.64		
				$R_{1.234}$	.65		

The coefficient of multiple correlation between female labor-force participation and the two factors, male earnings and percentage of women engaged in manufacturing and domestic service, was .64. The negative relationship between male earnings and percentage of women engaged in manufacturing and domestic service (−.42) thus raised the coefficient from .54 to .64. The addition of the fourth factor, percentage of the population white, did not change the coefficient of multiple correlation significantly.

INDUSTRIAL STRUCTURE AND FEMALE LABOR-FORCE PARTICIPATION

Further insight into the nature of the relationship between employment opportunities for women and the participation of women in the labor force may be gained by a qualitative analysis of the position of the various cities on the accompanying scatter diagram (Figure 1). This chart illustrates the relationship between the median earnings of males in the labor force in 1940 who worked 12 months in 1939, and the percentage of females 14 years of age and over in the labor force in 1940 in 91 cities over 100,000 population. The coefficient of correlation was −.54. The dotted lines indicate the averages of the two distributions. The cities outside the rectangle in the center, and particularly those in the second and fourth quadrants, are those which give the negative slope to the relationship. It is also mainly these cities which have industrial characteristics supporting the thesis that employment opportunities for women are a major factor in determining the proportion of women which engages in the labor force.

Below the average in proportion of women in the labor force are found the following cities (numbers refer to key in Table 2):

City	*Dominant Industry*
Gary (20)	Iron and steel
Youngstown (67)	
Erie (71)	
Pittsburgh (73)	
Buffalo (53)	

FIGURE 1
RELATIONSHIP BETWEEN MALE EARNINGS AND FEMALE PARTICIPATION IN THE LABOR FORCE, NINETY-ONE LARGE CITIES IN THE UNITED STATES: 1940.

City	*Dominant History*
Detroit (37)	Automobile
Flint (38)	
Fort Wayne (19)	
South Bend (22)	
Akron (60)	Basic rubber
Tacoma (90)	Lumber
Long Beach (2)	Petroleum and navy installation
San Diego (6)	Navy installation
Grand Rapids (39)	Furniture
Scranton (75)	Mining
Kansas City, Kan. (24)	None
Seattle (88)	
Salt Lake City (85)	
Duluth (40)	
Oakland (4)	
San Antonio (84)	See discussion below
Yonkers (58)	

Above the average in the participation of women in the labor force are:

City	*Employment opportunities for women in:*
Fall River (31)	Textiles and apparel
New Bedford (33)	Textiles
Reading (74)	
Lowell (32)	
Knoxville (78)	
Charlotte (59)	Textiles and domestic service
Chattanooga (77)	
Tampa (15)	Tobacco and domestic service
Richmond (87)	
Atlanta (16)	Domestic service
Dallas (81)	
Miami (14)	
Nashville (80)	
Memphis (79)	
Jacksonville (13)	
New Orleans (27)	
Providence (76)	State government, some textiles or apparel, and tobacco
Trenton (51)	
New Haven (11)	Apparel
Bridgeport (9)	Apparel and electrical machinery
Sacramento (5)	State government
Albany (52)	
Cambridge (30)	See discussion below
Hartford (10)	
San Francisco (7)	

It will be noted that in Yonkers, with high average earnings, women nevertheless engage in the labor force in average numbers where opportunities are afforded in textiles and apparel. In San Antonio, with extremely low earnings, women have not been pushed into employment, since there are few women's jobs in that community. Hartford and San Francisco, while without the light manufacturing industries of most of the other cities with similar female labor force propensities, probably provide an unusually large number of clerical jobs to women, particularly in the insurance companies located there.

It is granted that not all of the variation in the propensity of women to engage in the labor force is explained by the above analysis. Cambridge (30), for example, which ranks above Bridgeport (9) and Lowell (32) in the participation of women, does not appear to have comparable opportunities for the employment of women. There are other cities which, on the basis of their industrial structure, one would expect to find located in more extreme positions with respect to the participation of women in the labor force. Birmingham (1) and Norfolk (86), with large non-white populations, might be expected to rank with other southern cities which have large numbers of women engaged in domestic service.[7] The explanation may lie in the fact that the dominant industries of these areas are iron and steel, and a navy installation. Other cities with similar employment for men had even lower participation of women in the labor force. Baltimore (28), Fort Worth (82), Houston (83), and Louisville (26), might also be expected to follow the pattern of higher participation of women through employment in domestic service. Philadelphia (72), Camden (46), Elizabeth (47), and Paterson (50), with opportunities for women in light manufacturing industries, are nevertheless near the average in the proportion of women employed. On the other hand, Cleveland (63), Dayton (65), Milwaukee (91), and Worcester (36), with no outstanding industries to give employment to women, might be expected to have lower participation.

The negative relationship between male earnings and percentage of women engaged in manufacturing and domestic service, previously noted, is also apparent in these lists since the cities in which the dominant industries are extractive or durable goods have higher earnings than those in which the light industries predominate.

---

[7] All other cities with more than 20 per cent non-white population had female labor force participation rates above the 91-city average. The coefficient of correlation between the percentage of the population white and the percentage of females employed in domestic service was −.87.

CONCLUSION

There is strong evidence that the industrial structure of the community affects the propensity of women to engage in the labor force. This study has found that in the large cities of the United States in 1940, in general, high labor force propensities among women were found in areas in which either or both of two conditions existed: the concentration of an industry (textiles, apparel, electrical machinery, or tobacco manufacturing) in which jobs adapted to women are found in large numbers; or a social tradition (usually accompanied by a large Negro population) accustomed to the employment of large numbers of domestic servants. Conversely, in general, low labor force propensities among women were found in areas in which the dominant industries (iron and steel, machinery, automobile, and basic rubber manufacturing) have relatively few jobs which can be performed by women. Is it, then, the low level of earnings which *pushes* women into search for employment, or is it the presence of industries with jobs for women which *pulls* them into the labor force?[8]

At or near the subsistence level, low earnings undoubtedly force additional workers into the labor market. However, the calculation of coefficients of elasticity assuming a causal relationship at all income levels would appear to ignore the important factor of available job opportunities.

[8] A corollary question beyond the scope of this paper is: Are levels of earnings low because many of the jobs in some industries can be filled by women?

ESTIMATING PARAMETERS OF PEARSON TYPE III POPULATIONS FROM TRUNCATED SAMPLES*

A. C. COHEN, JR.
The University of Georgia

The method of moments is employed with "single" truncated random samples; (1) to estimate the mean, μ, and the standard deviation, σ, of a Pearson Type III population when α_3 is known; and (2) to estimate μ, σ, and α_3 when only the form of the distribution is known in advance. No information is assumed to be available about the number of variates in the omitted portion of the sample. The results obtained can be readily applied to practical problems with the aid of "Salvosa's Tables of Pearson's Type III Function." An illustrative example is included in the paper.

I. INTRODUCTION

METHODS for estimating the mean and standard deviation of normally distributed populations from single truncated random samples have previously been considered by Pearson and Lee [1, 2], Fisher [3], and this writer [4]. In the present paper, the method of moments is employed to develop formulas for estimating population parameters from single truncated samples when the population is distributed according to Pearson's Type III Function. Throughout this paper the population mean, standard deviation and third standard moment are designated in turn as μ, σ, α_3, and their respective estimates as m, σ', and α_3'.

II. EQUATION OF THE TRUNCATED DISTRIBUTION

In the range dx', the frequency of a Pearson Type III distribution with parameters μ, σ, and α_3, from which the left "tail" is truncated at x_0', may be written as

$$(1) \qquad f(x')dx' = \frac{C}{I_0(\xi)}\left[1 + \frac{\alpha_3}{2}\left(\frac{x' - \mu}{\sigma}\right)\right]^{4/\alpha_3^2 - 1} e^{-2(x'-\mu)/\sigma\alpha_3}\frac{dx'}{\sigma};$$

where $x_0' \leqq x' \leqq \infty$; ξ is the truncation point in standard units of the complete distribution; $I_0(\xi)$ is the area under a Type III curve with third standard moment α_3, which lies to the right of ξ; and

$$C = \left(\frac{4}{\alpha_3^2}\right)^{4/\alpha_3^2 - 1/2} e^{-4/\alpha_3^2}\left[\Gamma\left(\frac{4}{\alpha_3^2}\right)\right]^{-1}.$$

* Presented before a joint session of the American Statistical Association, the Institute of Mathematical Statistics, and the Econometric Society, in New York, New York, on December 29, 1949.

Hence we have

$$I_0(\xi) = \int_{\xi}^{\infty} f(t)dt, \tag{2}$$

and

$$\xi = \frac{x_0' - \mu}{\sigma}. \tag{3}$$

To simplify our work, we translate the origin to x_0' and designate the new variable by omitting primes ($x = x' - x_0'$). Equation (1) thus becomes

$$f(x)dx = \frac{C}{I_0(\xi)}\left[1 + \frac{\alpha_3}{2}\left(\frac{x}{\sigma} + \xi\right)\right]^{4/\alpha_3^2 - 1} e^{-2(x/\sigma + \xi)/\alpha_3}\frac{dx}{\sigma}; 0 \leqq x \leqq \infty. \tag{4}$$

III. MOMENT RECURRENCE EQUATION

In order to derive a moment recurrence equation for the truncated Type III distribution, we begin with the Pearson differential equation

$$\frac{1}{f(x)}\frac{df(x)}{dx} = -\frac{\sigma\left(\frac{\alpha_3}{2} + \xi\right) + x}{\sigma^2\left(1 + \xi\frac{\alpha_3}{2}\right) + \sigma\frac{\alpha_3}{2}x}, \tag{5}$$

which has for a solution, the Type III frequency function with origin at x_0'. If we separate the variables, multiply both sides of the resulting equation by x^k, and integrate over the truncated range $0 \leqq x \leqq \infty$, we obtain

$$\sigma\int_0^{\infty}\left[\sigma\left(1 + \xi\frac{\alpha_3}{2}\right) + \frac{\alpha_3}{2}x\right]x^k d(fx) = -\int_0^{\infty}\left[\sigma\left(\frac{\alpha_3}{2} + \xi\right) + x\right]x^k f(x)dx. \tag{6}$$

The right member of this equation becomes

$$-\sigma\left(\frac{\alpha_3}{2} + \xi\right)\mu_k' - \mu'_{k+1}; \quad k \geqq 0,$$

where μ_k' is the kth moment of the incomplete distribution about the point of truncation; i.e.

$$\mu_k' = \int_0^\infty x^k f(x)dx.$$

Upon integrating by parts, the left member of (6) becomes

$$\sigma\left\{\sigma\left(1 + \xi\,\frac{\alpha_3}{2}\right) + \frac{\alpha_3}{2}\,x\right\}x^k f(x)\Big|_0^\infty - k\sigma^2\left(1 + \xi\,\frac{\alpha_3}{2}\right)\mu'_{k-1} - (k+1)\sigma\,\frac{\alpha_3}{2}\,\mu_k'.$$

The first term in the above expression vanishes as x approaches the upper limit for all $k \geqq 0$. It vanishes at the lower limit for $k \geqq 1$, but not for $k=0$. When the left and right members above are combined we have

$$\sigma\left(k\,\frac{\alpha_3}{2} - \xi\right)\mu_k' + k\sigma^2\left(1 + \xi\,\frac{\alpha_3}{2}\right)\mu'_{k-1} = \mu'_{k+1}; \quad k \geqq 1. \tag{7}$$

Although Equation (7) does not hold when $k=0$, we may substitute this value in (6) prior to integrating and obtain

$$-\sigma\xi + \sigma\left(1 + \xi\,\frac{\alpha_3}{2}\right)\cdot\sigma f(0) = \mu_1'.$$

If we write

$$Z = \sigma f(0) = f(\xi)/I_0(\xi),$$

this result becomes

$$\sigma\left[Z\left(1 + \xi\,\frac{\alpha_3}{2}\right) - \xi\right] = \mu_1'. \tag{8}$$

Craig [5] previously obtained results similar to Equations (7) and (8), with moments referred to the mean of the complete distribution. The moments in (7) and (8) are referred to the point of truncation.

IV. ESTIMATES OF σ AND ξ WHEN α_3 IS KNOWN

Let $k=1$ in Equation (7) and we obtain

$$\sigma\left(\frac{\alpha_3}{2} - \xi\right)\mu_1' + \sigma^2\left(1 + \xi\,\frac{\alpha_3}{2}\right) = \mu_2'. \tag{9}$$

If we let

$$\theta = 1 + \xi \frac{\alpha_3}{2},$$

and substitute this value in (8) and (9), these equations can be written as

$$(10) \qquad \begin{cases} \mu_1' = \sigma[Z\theta - \xi], \\ \mu_2' = \sigma^2\left[\theta + \left(\frac{\alpha_3}{2} - \xi\right)(Z\theta - \xi)\right]. \end{cases}$$

Eliminate σ between the two equations above and we have

$$(11) \qquad \frac{\mu_2'}{(\mu_1')^2} = \frac{1}{Z\theta - \xi}\left[\frac{\theta}{Z\theta - \xi} + \frac{\alpha_3}{2} - \xi\right].$$

Since α_3 is assumed to be known in the present case, the right side of (11) which we subsequently designate as $Q(\xi)$, is here considered as a function of ξ alone. It is to be pointed out that results equivalent to (11), previously obtained [4] for the normal distribution, also follow from Equation (11) if we consider the normal distribution as a special case of the Type III distribution with $\alpha_3 = 0$, and hence $\theta = 1$.

Applying the method of moments, we equate truncated sample moments to corresponding population moments, and Equation (11) becomes

$$(12) \qquad \frac{\nu_2'}{(\nu_1')^2} - Q(\xi) = 0.$$

$Q(\xi)$ is defined by the right side of (11) and ν_k' is the kth moment of the truncated sample referred to the point of truncation; i.e.

$$\nu_k' = \frac{1}{n} \sum_1^n x_i^k,$$

where n is the number of sample observations.

Equation (12) can be solved by the method of successive approximations illustrated in Section VI, to determine ξ' which is of course an estimate of the population parameter ξ. $Q(\xi)$ for such values of ξ as may be necessary to complete the solution, can be readily evaluated with the aid of Salvosa's Tables of Pearson's Type III Function [6].

With ξ' determined from (12), the first equation of (10) now enables us to write the estimating equation

$$\sigma' = \frac{\nu_1'}{Z\theta - \xi'} \cdot \tag{13}$$

With ξ' and σ' determined from (12) and (13), an estimate of the population mean can be obtained from

$$m = x_0' - \sigma'\xi', \tag{14}$$

which follows from Equation (3).

V. A FIRST APPROXIMATION TO ξ'

Figure 1 contains a set of graphs of the function $Q(\xi)$ over the range $-3.0 \leqq \xi \leqq 3.0$, for α_3 from 0 to 1.0 at intervals of 0.2, which permits us to obtain a satisfactory first approximation to ξ' for use in solving Equation (12). Even without these graphs, a reasonably satisfactory first approximation might be obtained from the relation

$$\xi' \sim \frac{x_0' - \nu_1'}{\sqrt{\nu_2' - (\nu_1')^2}},$$

or perhaps by judicious guessing. If the first approximation differs by an appreciable amount from the true value of ξ', it will only be necessary to carry out one or two additional computations of $Q(\xi)$ above those normally involved, to obtain the required accuracy in the final approximation.

VI. A NUMERICAL EXAMPLE

The method of successive approximations employed in solving Equation (12) for ξ is the same as that previously used by the author [4] in solving similar equations which arise in connection with truncated normal samples. Only simple linear interpolation in conjunction with the Salvosa Tables of the Pearson Type III Functions is involved. The technique can best be explained by considering the following numerical example in which $\alpha_3 = 0.6$; $n = 614$; $x_0' = 110.5$; $\nu_1' = 32.6596$; and $\nu_2' = 1359.2793$. This example was constructed by arbitrarily truncating a sample distribution of the Weights of 629 University of Washington Freshmen so that all record of observations less than 110.5 lbs. is assumed to be missing. Fifteen variates are in the omitted portion of the distribution. The class interval of this distribution is 11 lbs., and there are ten such classes in the retained portion. No grouping corrections were applied in computing the moments for the truncated sample. Using the moment values given above we find that

$$\nu_2'/(\nu_1')^2 = 1.27434.$$

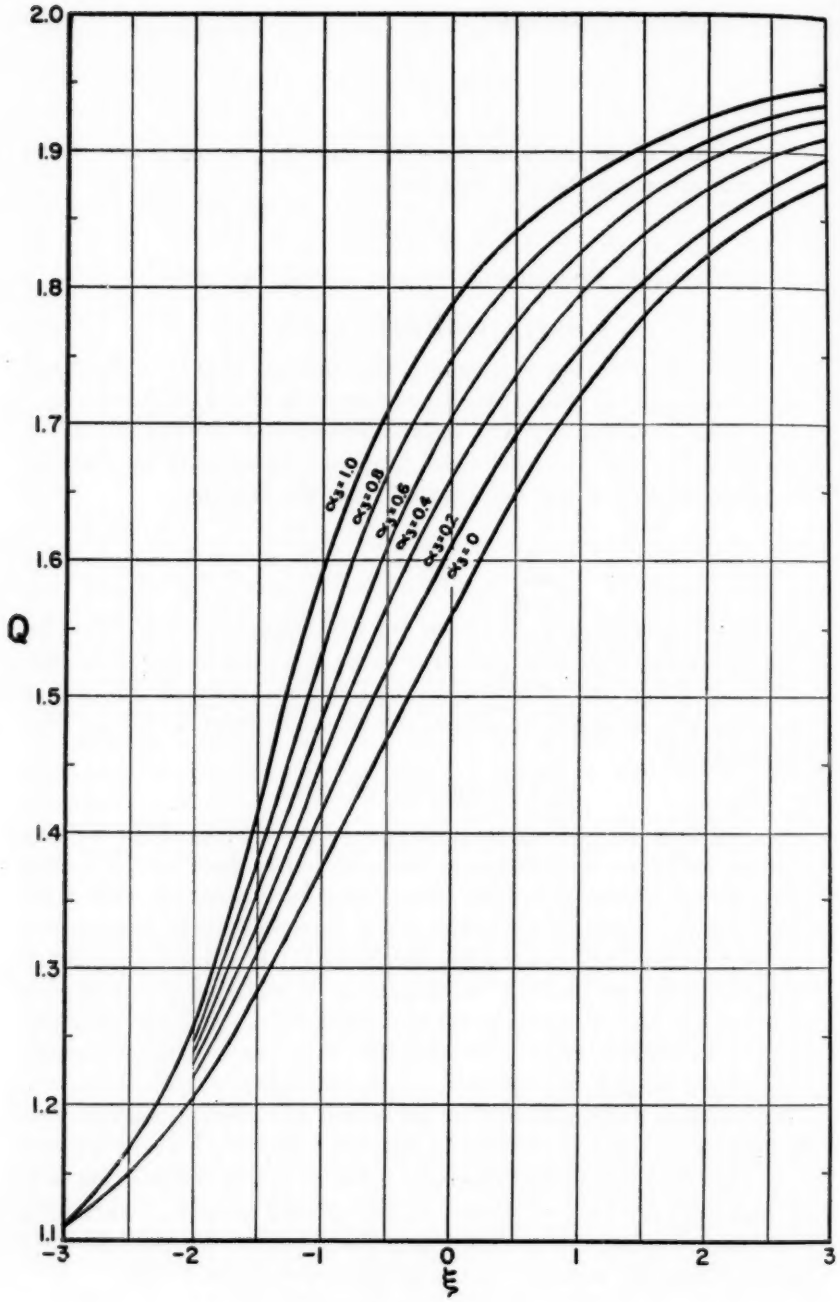


FIGURE 1

GRAPHS OF THE FUNCTION Q

To find ξ' we must then solve the equation

$$Q(\xi) = 1.27434,$$

where $Q(\xi)$ is defined by the right side of Equation (11).

As a first approximation to ξ' we obtain the value $\xi_1' = -1.8$ from the $\alpha_3 = .6$ curve of Figure 1. By resorting to Salvosa's Tables we obtain $Q(-1.8) = 1.28567$, which being greater than 1.27434, indicates that a value less than -1.8 should be selected as the next approximation to ξ'. Accordingly, we select -1.9, and find that the value sought lies between these two approximations. Linear interpolation subsequently leads us to determine that $-1.85 < \xi' < -1.84$. On interpolating between these two values we obtain as a final approximation, $\xi' = -1.850$ which is correct to the nearest third decimal. Details of the computations involved above will be found in Table 1 which follows.

TABLE 1

COMPUTATIONS OF $Q(\xi)$

$\alpha_3 = 0.6$

ξ	$I_3(\xi)$	$f(\xi)$	$Z = \frac{f(\xi)}{I_3(\xi)}$	$\frac{1}{Z\theta - \xi}$	$\frac{\theta}{Z\theta - \xi} + \frac{\alpha_3}{2} - \xi$	$Q(\xi)$ (5) · (6)
1	2	3	4	5	6	7
−1.8	.985628	.062162	.063068	.546744	2.351502	1.28567
−1.9	.990897	.043868	.044271	.521095	2.424071	1.26317
−1.84	.987957	.054370	.055033	.536292	2.380259	1.27651
−1.85	.988491	.052521	.053132	.533719	2.387505	1.27426

From the fifth column of Table 1 we find that

$$\left.\frac{1}{Z\theta - \xi'}\right|_{\xi' = -1.850} = 0.533719.$$

Therefore from Equation (13) we have

$$\sigma' = (32.596)(.533719) = 17.43.$$

From Equation (14) we now find

$$m = 110.5 - (17.43)(-1.850) = 142.74.$$

The estimates $m = 142.74$ and $\sigma' = 17.43$ as obtained above, are to be compared with the corresponding values 142.25 and 17.59 which were

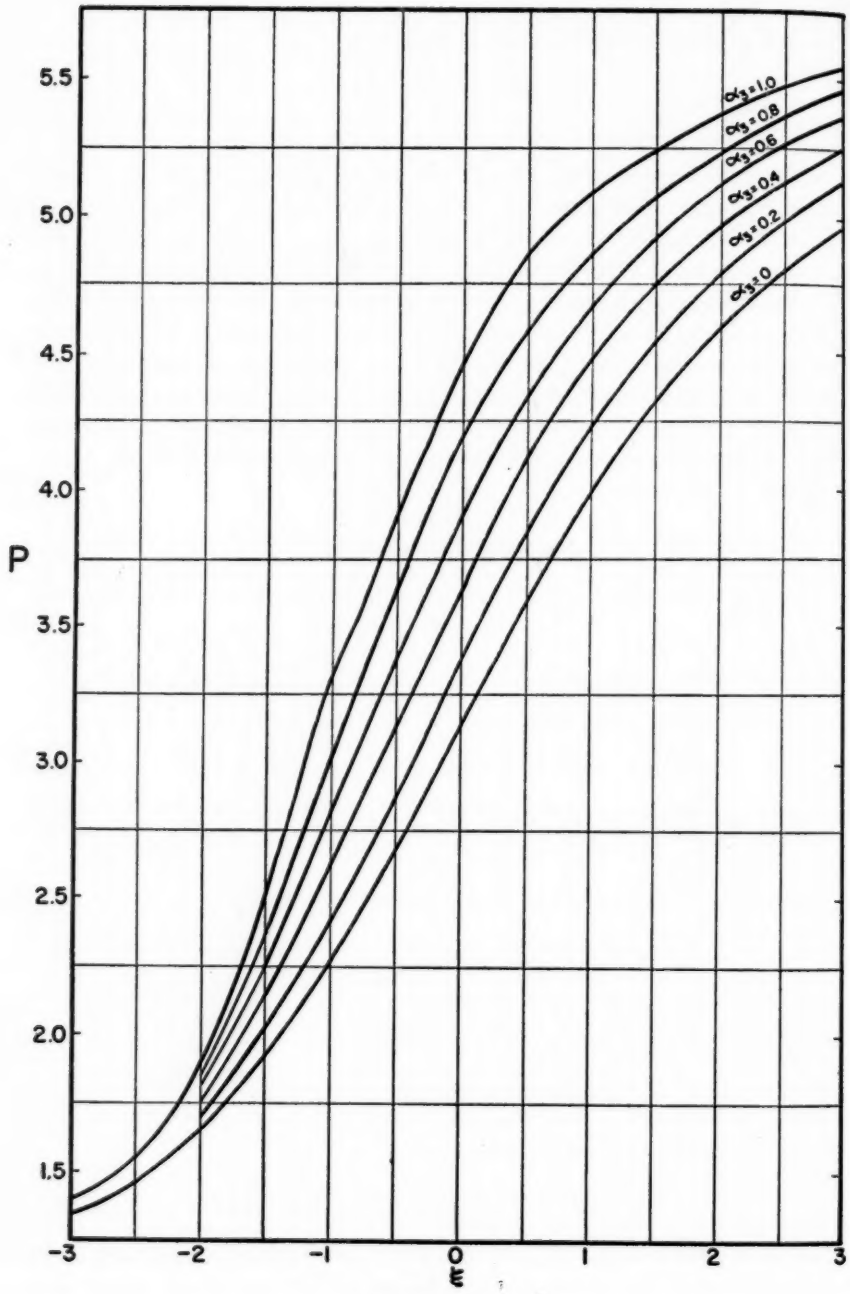


FIGURE 2

GRAPHS OF THE FUNCTION *P*

determined from the complete sample. The discrepancies between the two sets of values may be attributed in part to the fact that the observed distribution is not a true Type III distribution even though the universe sampled might have been of this type. Furthermore, no grouping, corrections were made in computing moments of the truncated sample. The usual Sheppard's corrections are not applicable in this case because of the abrupt termination of the curve at the point of truncation. The Pairman [7] corrections which are more tedious to apply, however, could have been used.

VII. CASE WHEN α_3 IS NOT GIVEN

When α_3 is not known in advance and must be estimated from the truncated samples along with the mean and standard deviation, the situation becomes somewhat more complex. In this instance we require an additional moment equation involving the third moment of the truncated distribution. Let $k=2$ in Equation (7) and simplify by substituting values of μ_1', and μ_2' from (10) to obtain

$$(15)\quad \mu_3' = \sigma^3\left\{\theta(\alpha_3 - \xi) + (Z\theta - \xi)\left[\left(\frac{\alpha_3}{2} - \xi\right)(\alpha_3 - \xi) + 2\theta\right]\right\}.$$

Eliminate σ between Equations (8) and (15) to obtain

$$(16)\quad \frac{\mu_3'}{(\mu_1')^3} = \frac{1}{(Z\theta - \xi)^2}\left[\frac{\theta(\alpha_3 - \xi)}{Z\theta - \xi} + \frac{\alpha_3}{2}(\alpha_3 - \xi) + \xi^2 + 2\right].$$

Again equating sample moments to corresponding population moments we have

$$(17)\quad \frac{\nu_3'}{(\nu_1')^3} - P(\xi, \alpha_3) = 0,$$

where $P(\xi, \alpha_3)$ is defined by the right side of (16), and ν_1', and ν_3' are truncated sample moments as previously defined. To determine the estimates ξ' and α_3', we must simultaneously solve Equations (12) and (17) with Q in Equation (12) now being considered as a joint function of both ξ and α_3. For clarity we will subsequently write $Q(\xi, \alpha_3)$ to emphasize this fact. After ξ' and α_3' are determined, estimates of the mean and standard deviation may be obtained as before from Equations (13) and (14). To facilitate the simultaneous solution of (12) and (17), a set of graphs of $P(\xi, \alpha_3)$ over the range $-3.0 \leqq \xi \leqq 3.0$ for α_3 from 0 to 1.0 at intervals of 0.2, is presented in Figure 2.

VIII. SOLUTION OF SIMULTANEOUS EQUATIONS

As a practical means of carrying out the simultaneous solution of Equations (12) and (17), a modified Newton-Raphson method which is described by Whittaker and Robinson[1] has been employed by the writer. Satisfactory first approximations to ξ' and α_3' can be read directly from the graphs in Figures 1 and 2. The necessary areas and ordinates required for computing various values of $P(\xi, \alpha_3)$ and $Q(\xi, \alpha_3)$ can be found in Salvosa's Tables. To illustrate this procedure we again use the sample data given in Section VI. In this instance, however, we consider α_3 as not being known, and we therefore estimate its value from the sample. From the given data we compute $\nu_3' = 68{,}269.15$ and

$$\nu_3'/(\nu_1')^3 = 1.93100.$$

To determine ξ' and α_3' we must simultaneously solve

$$Q(\xi, \alpha_3) = 1.27434,$$
$$P(\xi, \alpha_3) = 1.93100,$$

where Q and P are defined by the right sides of Equations (11) and (16) respectively.

From the curves of Figures 1 and 2 we obtain as first approximations, $\xi' = -1.9$, and $\alpha_3' = 0.7$. Holding α_3 constant, we solve $Q(\xi, \alpha_3 = 0.7) = 1.27434$ for ξ in the manner of Section VI to obtain -1.875. In the same way we solve $P(\xi, \alpha_3 = 0.7) = 1.93100$ for ξ and obtain -1.874. When we let $\alpha_3 = 0.8$ and solve $Q(\xi, \alpha_3 = 0.8) = 1.27434$, we obtain $\xi = -1.894$. Similarly, on solving $P(\xi, \alpha_3 = 0.8) = 1.93100$, we obtain $\xi = -1.900$. Linear interpolation between these values, yields the estimates

$$\xi' = -1.878 \quad \text{and} \quad \alpha_3' = 0.71.$$

If necessary, the above process could be repeated to improve the accuracy with which estimates are computed. Detailed computations involved in the various steps are set forth in Tables 2 and 3 which follow.

With ξ' and α_3' thus determined, σ' and m are computed as in Section VI. Interpolation from Table 2 gives

$$\left.\frac{1}{Z\theta - \xi}\right|_{\substack{\alpha_3 = 0.71 \\ \xi = -1.878}} = 0.529,$$

[1] Whittaker, E. T. and Robinson, G., "The Calculus of Observations," Blackie and Sons, Ltd., London and Glasgow, Second Edition, 1929, par. 46, page 88.

TABLE 2

COMPUTATIONS OF $P(\xi, \alpha_3)$ AND $Q(\xi, \alpha_3)$

ξ	α_3	$I_s(\xi)$	$f(\xi)$	Z	$\frac{1}{Z\theta-\xi}$	$Q(\xi, \alpha_3)$	$P(\xi, \alpha_3)$
−1.85	.7	.992136	.044485	.044848	.535962	1.28037	1.95240
−1.90	.7	.994134	.035633	.035843	.523011	1.26842	1.90810
−1.85	.8	.995474	.034068	.034223	.537953	1.28464	1.98325
−1.90	.8	.996953	.025359	.025436	.524630	1.27271	1.93144
−1.91	.8	.997199	.023805	.023872	.522020	1.27018	1.90879

TABLE 3

DIFFERENCES OF ξ AND α_3

α_3	ξ from $Q(\xi, \alpha_3)$	ξ from $P(\xi, \alpha_3)$	Difference
.7	−1.875	−1.874	−.001
.8	−1.894	−1.900	+.006

and from Equation (13) we have

$$\sigma' = (32.596)(0.529) = 17.25.$$

From Equation (14) we find

$$m = 110.5 - (17.25)(-1.878) = 142.88.$$

The estimates $m = 142.88$, $\sigma' = 17.25$, and $\alpha_3' = 0.71$ are to be compared with corresponding complete sample values 142.25, 17.59, and 0.59. With the exception of the estimates of α_3, the agreement between corresponding values is quite close. The disparity between the two estimates of α_3 while due in part to the same reasons already given in explaining discrepancies noted in Section VI, may be largely attributed to the fact that the omitted portion of the sample does not conform to the shape of the distribution as determined by the retained portion. A graduation performed on the truncated distribution using the values 142.88, 17.25 and 0.71 for m, σ' and α_3' respectively, fitted the observed data much better than did a graduation using corresponding estimates obtained from the complete sample. The only appreciable discrepancy between observed and graduated cell frequencies occurred in the omitted (first) cell. This result suggests that perhaps this cell contained a number of misplaced observations which caused an appreciable error in computing the third moment of the complete sample and consequently an error in α_3'.

IX. PRECISION OF ESTIMATES

A careful investigation of the precision of the estimates obtained in this paper remains to be carried out. However, the following well known formulas for the variances of sample moments for large samples may be useful in the absence of more exact results:

$$(18)\qquad \begin{cases} V(\nu_1') = \dfrac{\mu_2}{n}, \\ V(\nu_2) = \dfrac{\mu_4 - \mu_2^2}{n}, \\ V(\nu_3) = \dfrac{\mu_6 - \mu_3^2 - 6\mu_4\mu_2 + 9\mu_2^3}{n}, \end{cases}$$

where the ν's are sample moments and the μ's are population moments. The omission of primes designates central moments, which in this case are referred to the mean of the truncated rather than complete distribution.

By using the recursion formula of Equation (7), the above variances can be expressed in terms of σ, α_3 and Z only; for example,

$$(19)\qquad V(\nu_1') = \frac{\sigma^2}{n}\left[\theta - (Z\theta - \xi)\left(Z\theta - \frac{\alpha_3}{2}\right)\right].$$

Similar expressions for the variances of ν_2 and ν_3 can also be written, but they are considerably more involved.

Further, it seems reasonable to suppose that for small values of α_3 such that the distributions are in fact approximately normal, that the variances of ξ' and σ' will be approximately equal to corresponding variances previously obtained [3], [4] for truncated normal distributions. When the number of sample observations is large enough, the χ^2 test may appropriately be employed to determine whether a fitted curve is in satisfactory agreement with observed data.

REFERENCES

[1] Pearson, Karl, and Lee, Alice, "On the Generalized Probable Error in Multiple Normal Correlation," *Biometrika*, Vol. 6, 1908, pp. 59–68.

[2] Lee, Alice, "Table of Gaussian 'Tail' Functions when the 'Tail' is Larger Than the Body," *Biometrika*, Vol. 10, 1915, pp. 208–215.

[3] Fisher, R. A., "Properties and Applications of *Hh* Functions," Introduction to *Mathematical Tables*, Vol. 1, British Association for the Advancement of Science, 1931, pp. xxvi–xxxv.

[4] Cohen, A. C., Jr., "On Estimating the Mean and Standard Deviation of Truncated Normal Distributions," *Journal, American Statistical Association*, Vol. 44, No. 248, December 1949, pp. 518–525.

[5] Craig, C. C., "The Semi-Invariants and Moments of Incomplete Normal and Type III Frequency Functions," *Annals of Mathematics*, Second Series, Vol. 31, No. 2, May 1930, pp. 251–270.

[6] Salvosa, Luis R., "Tables of Pearson's Type III Function," *Annals of Mathematical Statistics*, Vol. 1, May, 1930, appended.

[7] Pearson, Karl and Pairman, Eleanor, "On Corrections for the Moment Coefficients of Limited Range Frequency Distributions when There Are Finite or Infinite Ordinates and Any Slope at the Terminals of the Range," *Biometrika*, Vol. 12, pp. 231–258.

RESPONSE ERRORS IN CENSUS RESEARCH

ELI S. MARKS
AND
W. PARKER MAULDIN
Bureau of the Census

Some efforts of the Bureau of the Census to measure and evaluate response errors are described. Particular attention is devoted to the October, 1948, pretest of Census procedures and the measurement of response errors of the various procedures. The procedures were: (1) Self-enumeration, using a household schedule. The schedule was a single sheet of paper 18″×24″ with 10 columns of questions for the various members of the household—two columns of questions for each individual repeated five times—on one side, and a letter to the household and questions on housing on the other side. (2) Self-enumeration using a 4 page booklet, the page size being 9″×11″, with Individual Enumeration Forms (9″×11″) inserted for each member of the household. The questions were printed front and back in two columns on the Individual Enumeration Form. (3) Obtaining information from respondents by direct interview. (4) Obtaining basic information by direct interview and leaving schedule with additional questions for respondent to fill out, the schedule to be picked up on a second visit to the household.

Approximately the same proportion of respondents filled the two self-enumeration forms, and the error rates on the two forms also were approximately the same. Furthermore, the total enumerator time per household was almost exactly the same for the two self-enumeration procedures. Direct field costs of self-enumeration did not exceed those for direct enumeration in spite of the higher callback rate. Also, the leave-it procedure was not more expensive than direct enumeration even in rural enumeration districts where the cost of callbacks is particularly high.

IN THE COURSE of a fairly long history, the Bureau of the Census has devoted a large amount of time to research designed to improve the quality of data taken in censuses or designed to reduce the cost of taking the census. This research has involved problems of classification and of sampling. More recently we have embarked upon a program aimed at the problems of response errors in censuses and in sample surveys. Among the earliest work along these lines was the investigation of responses in the Monthly Report of the Labor Force. The quality check conducted as a part of the 1945 Census of Agriculture

was the first measurement of response errors incorporated in a regular census. The 1947 Census of Manufactures involved rather extensive experimentation with question wordings and other phases of design. It also included a coverage check. The 1948 Census of Business was followed by a Post-Enumeration Survey designed to measure the accuracy of the census. The present paper is based on experiments conducted in connection with the three major censuses taken in 1950 (i.e., the Censuses of Population, Housing and Agriculture).

There are three general types of error possible in a survey—sample, response, and processing. By response error we mean the error arising in the course of obtaining and recording information, whether the error is due to the respondent, the interviewer, or to the design of the interview.

Survey errors may be classified in various ways. A classification and discussion of such errors appears in a paper by Deming [1]. In the discussion that follows we shall be talking primarily about errors that arise from the design of the interview, the wording of the questions, the procedure followed by the interviewer, and the conscious and accidental errors on the part of the respondent. It will be noted that the research described in this paper is primarily concerned with the measurement of total response error, rather than with the analysis of response error components.

Three general comments on the direction of the Bureau's response research program are in order: First, the philosophy underlying our sampling research activities—devise the most efficient plan for the money—has been carried over to this newer program. The necessity for maintaining complete cost records, even for those alternative procedures for which seemingly reasonable *a priori* estimates of relative cost have been made, has been pointed up in our research experience. On occasion the alternative which has been thought almost too expensive to test has turned out to be the less expensive.

The investigation of *coverage* has been made an important aspect of the response research program. One of the most important tasks in census-taking is to obtain an accurate count, whether of people, or of dwelling units, or of land in farms or of sales of business establishments. The task of getting "coverage" affects schedule design and question-wording as well as the setting up of procedures. An error which leads either to under- or over-enumeration of the population is as much a response error as is a misclassification of labor force status or a misreporting of age.

In devising procedures for the 1950 Census, the Bureau has to keep

constantly in mind this problem of large-scale control and the fact that a technique or question which looks perfect in Washington may be completely unsatisfactory in Dothan, Alabama. The lesson has been hard to learn (and reasonably expensive) but we are gradually becoming convinced that experience and judgment (no matter how good they be) can never substitute for actual field testing. Part of this lesson has been the pretests for the 1950 Census. We shall concentrate on one of these tests—the October 1948 Pretest.

THE OCTOBER 1948 PRETEST

A complete description of this test would be fairly lengthy. We shall, therefore, merely list the main purposes of the October Pretest. These were:

1) To determine whether gains in either cost or accuracy could be made by using a self-enumeration procedure, i.e., a procedure in which respondents filled out the schedules. Population and Agriculture Census procedure in the United States has involved an enumerator interviewing the respondent and recording all the answers. Other countries, however, have used a self-enumeration procedure, some apparently with success.

2) To determine whether gains could be made in cost and accuracy by use of a short basic schedule which applies to everyone, supplemented by a longer schedule taken on a sample basis.

3) To determine the relative efficiency of distributing self-enumeration schedules by mail compared with a distribution by an enumerator who completes a very short schedule by direct enumeration and then leaves the complete Census schedule to be filled out by the respondent.

4) To determine the relative cost and accuracy of various schedule formats.

For each of the procedures used, a detailed Time and Call Record was kept of each activity so that it was possible after the enumeration was completed to tell how much time had been spent in actual enumeration, in travel between households, and in all other activities. As an absolute measure of cost, the Time and Call Records are subject to considerable error, but they are extremely useful in measuring cost differences between procedures.

To assess accuracy of the various schedules, a check was made of a substantial proportion of the households and farms, using questionnaires with more detailed questions about selected topics. Whenever the original entry differed from the answer obtained on the check interview, the respondent was asked to explain the discrepancy. The inter-

viewers used in the quality check were professional personnel from the Washington office who were well oriented toward the objectives of the survey. We are confident that the information obtained on the quality check is somewhat more accurate than data obtained on the original enumeration, although it is obvious that one may obtain two different answers without getting the correct answer.

The following results of the October Pretest are of particular interest:

1) The October Pretest involved a complete census of four counties: Carroll and Oldham Counties, Kentucky; Putnam County, Illinois; Union County, Indiana—and two census tracts in Minneapolis, Minnesota. In Oldham and Putnam counties the schedules were distributed by the Post Office. A letter was printed on the schedules asking the respondent to fill out the schedule. The schedules were later picked up by a census enumerator who completed those schedules not filled out by the respondent. The schedule used in Putnam differed considerably in format from that used in Oldham although the content and question wordings were identical.

In Putnam County the form used for enumerating Population and Housing was Schedule D. One schedule was used for each household containing less than six persons (for larger households an additional schedule was used as a continuation sheet). One side of this schedule contained a letter to the household, the questions on housing, and a section to be completed by the enumerator. The other side contained 10 columns of questions for the various members of the household—two columns of questions for each individual repeated five times. To complete this side of the schedule, the respondent had to spread it out to its full size (18″×24″) a rather formidable expanse of paper almost completely covered with fairly fine print. When it came off the printing press its own parents called it an atrocity. Very few of us expected it to survive its early childhood. However, the schedule has substantial advantages in processing, and, therefore, it seemed worth testing.

In Oldham County the form used for enumerating Population and Housing was Schedule A with the accompanying Individual Census Forms I. The A form was a 4 page booklet, the page size being 9″×11″. The first page contained the letter to the householder, the second page contained instructions and space for listing the members of the household, the third page contained questions on housing, and the fourth page carried the items to be filled in by the enumerator. For each member of the household, one of the Individual Enumeration Forms I was used. This was a sheet of paper 9″×11″. The questions were printed front and back in two columns. Although the print on the individual

form was smaller than is desirable, the schedule arrangement was reasonably "open" and did not appear to be nearly as crowded as did the questions on the D schedule. Most of the people concerned with the design of the schedules considered the A and I combination superior to the D schedule in appearance and convenience for the respondent. Certainly the A and I schedules were easier to handle and the respondent did not need the whole dining room table to spread them out.

Perhaps the respondents in Putnam County where the D schedule was used have big tables and hardy constitutions. A more likely hypothesis is that within fairly broad limits the basic tendency of respondents to cooperate with a government request for information is not appreciably affected by the schedule design. At any rate, about the same proportion of respondents (approximately $\frac{1}{2}$) filled out both forms, and the error rates on the two forms also were approximately the same. Furthermore, the total enumerator time per household for D is almost exactly the same as for A. Since only one county was used with each of the forms, the results cannot be conclusive. However, the lesson we have learned is that it is not safe to rely heavily upon your own judgment in deciding these matters.

2) In the other areas, one procedure used called for leaving a schedule at certain sample households and asking the household to fill it out. The enumerator returned a few days later to pick up the schedule. A parallel procedure used in the same areas called for the enumerator getting the same information by direct enumeration on his first call. Enumerators and enumerator assignments were allocated to one of the two procedures by a random process so that differences in excess of random variation between the results of the two procedures are, presumably, due to the procedure. The enumerators were aware that two procedures were being used and that they were working on a Pretest of the Census but the experimental aspect was not stressed in the enumerators' instructions. Several members of the Bureau staff felt that the "leave-it" procedure, since it involved an extra call, would mean such a large increase in travel time that its cost would be prohibitive. Serious doubts were cast on the wisdom of testing a procedure which was, it was assumed, much too expensive to be used. The results show that the callbacks required with the leave-it procedure do not make the procedure prohibitively expensive. In these tests, direct field costs of self-enumeration did not exceed those for direct enumeration in spite of the higher callback rate. Furthermore, the leave-it procedure was not more expensive than direct enumeration even in rural enumeration districts where the cost of callbacks is particularly high.

This experience suggests that methods should not be discarded just because someone thinks "they are too expensive to use." It is likely that many excellent techniques have been rejected on cost grounds when a test might have shown trivial cost differences.

3) Of course, not all of the results are in directions opposite to that expected. The items on the schedule which, prior to the test, were thought to be least accurate proved, in general, to be least accurate. "Highest grade of school completed" was the least accurate of the population items checked, with about 17 per cent of the original entries differing from the check entries. Age, on the other hand, was moderately well reported. For each of the procedures used, less than 5 per cent of the individuals were placed by the quality check interviewer in a different 5 year age group than the one reported by the original census enumerator. All of the figures are, of course, subject to sampling error. It should be noted that since all of the supervisory personnel in the Pretest were experienced employees of the Census Field Division, conditions were presumably more favorable from the quality standpoint than can be expected in a national census. To a certain extent, the quality check results may represent a "lower bound" of error. All we could measure were those errors which were *detected* by the check enumerators. It has also been suggested that, since the quality check interviewer had the results of the original enumeration, he might be influenced to record agreement with the original results in doubtful cases. While this is possible, evidence from a quality check conducted following a later Census Pretest (May 1949) tends to discount this possibility. In this test the quality check enumerators were given the original results only for every other household. A comparison of results where the enumerators had the original findings with those where they did not indicates that, for some items, there were fewer discrepancies when the check enumerator had the original findings than when he did not but that (a) for several items the differences were reversed and (b) none of the differences were statistically significant. This study was a small scale test involving only 114 households in all and these households were selected in a non-random fashion so that they can be taken as representing no population other than themselves. It is, then, entirely possible that significant differences might have been found if large representative samples of the entire population of the United States had been used.

4) The estimated differences in age reporting vary from procedure to procedure and from area to area. In Minneapolis, where two census tracts were covered, in 9.7 per cent of the cases the original and check

entries differed. Each of the other four areas covered had 4 per cent or less of the cases in which the original and check entries differed. This is, however, *gross* difference. Actually the *net* differences in age are fairly small, particularly when broad age groups are considered. For example, in Minneapolis, the original enumeration shows 38.1 per cent of the quality check sample cases in the age group 20–44 and 75.0 per cent aged 20 and over. Corresponding figures obtained in the check enumeration are 38.5 per cent and 75.0 per cent, or a net difference of only 0.4 per cent in the former case and no net difference in the latter. Furthermore, in all 10 year age groups the net differences are small, being less than 0.55 per cent for each area in the Pretest. It should be noted that these figures are net differences exclusive of errors introduced by over- or under-enumeration.

Thus, in the case of age, "compensating errors" did occur and the net error found was rather small. However, generalization of this fact to assume that *all* errors compensate is quite unwarranted. For example, the estimated net differences in the proportion living in the same county a year ago was as high as 2.0 per cent for one of the five areas. Similarly, the net difference in the proportion of persons finishing the 8th grade went as high as 2.8 per cent in one area and was over 1.3 per cent in all areas. Furthermore, even if *all* response errors compensate perfectly for a one dimensional variable, the results of cross-tabulations can still be badly distorted.

5) In the population and housing censuses we are, in general, asking for information which is known fairly exactly by the respondent in many cases. In the agriculture census, we are more frequently asking for information which is known by the respondent only approximately. This is true in particular for production and value items. Many farms may not keep any records for such data. When asked for this type of information the farmer may give, at best, a rough estimate based on memory and inference.

Unfortunately, the difference between "knowledge" and "estimate" may not always show up in a quality check. For example, experience would indicate that farmers can report acreages of hay harvested more accurately than tons of hay produced. However, in the quality check sample for two hay crops in each of the four counties the net differences for acreage were smaller than the net difference for tonnage for only three of the eight figures. While the estimated net differences were subject to large sampling errors, this lack of superiority in the acreage reports is not (statistically) consistent with our prior knowledge of the validity of the reporting of the two items.

Estimates of acreages harvested are usually consistent although there can be, and is, variation in reporting and recording such acreages. Responses on hay tonnage harvested are usually estimated on the basis of tons per acre, number of bales, or number of wagon loads. Bale weights are known only approximately, and wagon loads are guesses (e.g., "about a ton" or "there's the wagon, Mister. How much would you guess she holds?"). In spite of this, hay production figures are, in general, no worse than hay acreage figures when the quality check is used as a measure. It appears likely that this is a limitation of the quality check itself.

It is possible that information which is well known to and frequently reported by a respondent may show *less* consistency than information which the respondent has to supply by "guessing" and which he uses only rarely in his every day life. When asked for a well-known "fact" the respondent's attention is directed toward recall of the *fact*. When asked for information which is not readily available, the respondent's attention may be focused on recall of a previous report ("Let's see, what did I say last time?"). This difference in the direction of recall effort may well be reflected in a decreased variation from trial to trial in reporting less accurate material.

Another important source of "consistent error" is the tendency of respondents to use the same method of estimation on repeated interviews and even to use the same "estimating factor"—for example, corn production is frequently estimated by multiplying acres harvested by the farmer's guess as to the yield per acre. Further, these guesses tend to be remarkably consistent not only from interview to interview with the same farmer but even from farm to farm within the same area.

The above is an hypothesis for further investigation and is in no sense a conclusion which can be drawn from data now available. In quality checks for more recent Census Pretests, attention has been given to getting additional information on methods used by respondents in estimating production.

6) Thus far we have discussed differences indicated by our quality check. A major difficulty with a quality check through reinterviewing is that it does not detect consistent respondent error unless there is a substantial shift in approach. One major source of such consistency is the "conditioning" effect of the first interview. The results already reported for hay production reflect, in part, this conditioning effect. Another example is the question on the housing schedule which asked for an estimate of the amount for which owned dwelling units would rent by the month. In spite of a lapse of two weeks or more between

original enumeration and check enumeration, respondents would reply to this question "I don't really know. We've never rented this place. I told the lady who was here before $50 and that's about the best I can guess."

A different approach to measurement of error is the examination of internal evidence in the original reports. For farms in Putnam County which reported oats harvested for grain and soybeans harvested for beans, yields per acre were calculated by dividing acres harvested into bushels harvested. In 55 per cent of the cases for oats and 49 per cent for soybeans, the quotient was *exactly* an integer (e.g., exactly 32 bushels per acre). Furthermore, 29 per cent of the oat reports gave yields of 40, 50, or 60 bushels per acre. Twenty-eight per cent of the farms reporting soybeans show yields of exactly 20, 25, or 30 bushels per acre. It is a reasonable inference that, in these cases, production was obtained by multiplying acreage by estimated yield per acre. Furthermore, the high proportion of yields per acre of exactly 40, 50, or 60 bushels for oats and exactly 20, 25, and 30 bushels for soybeans suggests either rounding or rough estimates.

In addition to the 29 per cent of the oat reports for which this yield is exactly 40, 50, or 60 bushels per acre there are 22 per cent of the oat reports in which production as given ends in 00 and 10 per cent in which it ends in 000, i.e., reports which give production to nearest 100 or nearest 1,000 bushels. Thus about 60 per cent of the oats reports appear to show evidence of gross "rounding." The corresponding percentage for soybeans is 47 per cent. As a matter of fact, only 10 per cent of all the soybeans reports for Putnam County showed no evidence of rounding.

Over a large number of farms, rounding may have little or no effect on totals or averages. For small numbers of farms, rounding of this magnitude can be very serious.

7) The overall differences in accuracy between the different methods were too small and varied too much from area to area for definite conclusions to be drawn. However, there is evidence that some sources of response error are operating to a varying degree in the different methods. For instance, rounding off piles up responses unduly at certain points. In reporting highest grade completed[1] grades which represent completion of school tend to be over-reported: response errors are, therefore, especially large when completion of eighth grade, twelfth grade and senior year of college are reported in the original enumera-

[1] Only individuals who were not attending school at the time were counted in this analysis.

tion. This error is small when the respondent has had a chance to check back or to look at records: In the procedure where the interviewer completed the schedule on the first call, 19 per cent of the individuals who were reported as having completed the eighth grade, 12 per cent of those reported as having completed twelfth grade and 32 per cent (7 out of 22) of those reported as having completed college were placed by the quality check interviewer in a different category. The corresponding figures for the procedure where the enumerator left the schedule to be picked up are 17 per cent, 6 per cent, and 17 per cent (3 out of 18).

Similarly, for age, figures like 40 and 65 are convenient rounding off points. The quality check changed by one year or more 20 per cent of the individuals reported as 40 years old and 24 per cent of those reported 65 years old in the "take-on-first-call" procedure. In the "leave-it" procedure the corresponding percentages are 17 and 22. These percentages are, however, based on relatively few cases (between 23 and 49).

For some items ignorance of Census definitions on the part of the respondent may be an important source of error. Where this is true, it might be anticipated that direct enumeration would be more accurate than self-enumeration. One such item is the report on whether the individual "worked" last week. "Work" is defined to include all work for pay or profit and work in the operation of the farm, business or profession of another member of the family and to exclude housework and other work around the house. It is frequently difficult to get the respondent to understand the idea of including unpaid work on a family farm or in a family member's business or profession. In the October Pretest the ratios of persons reported as "working last week" in the original enumeration to persons reported in the quality check were:

	Male	*Female*
Direct enumeration procedure	.97	.92
Self-enumeration procedure	.99	.81

For males the difference in the (net) errors of the two procedures is very small. There is, however, a substantial difference in the net errors for females and the undercount (relative to the quality check results) is larger for the self-enumeration procedure. These results are consistent with our hypothesis since a large proportion of the persons originally reported as not working were unpaid family workers and this category is, in general, more important for women than for men (and also this type of work is more likely to be over-looked for women than for men).

The conditions under which one type of procedure produces better data than another certainly require further exploration. The results reported here are tentative and are applicable only to the particular population studied (which was not a representative sample of the United States population). The results apply only to the procedures used and it should be noted that schedules for over one-fourth of the households in the "self-enumeration" procedure were actually prepared almost entirely by the enumerator (because the householder failed to complete the schedule himself).

8) As already noted, the quality check showed only minor differences between the quality of data yielded by the various procedures. On the other hand, there were substantial differences in quality between counties. For example, on age the gross per cent differences were: Carroll, 3.4 per cent; Oldham, 3.0 per cent; Union, 1.8 per cent; Putnam, 1.5 per cent.

The relatively large differences in discrepancy rates between counties parallels other differences in county characteristics. It appears likely that an explanation of the phenomenon is to be found in the type of respondents and enumerators. One promising line of further research is investigation of the variation in error rates among classes of enumerators and classes of respondents.

9) With regards to coverage, net differences in number of farms between the quality check and the Census varied between counties from a low of 1.5 per cent to a high of 4.0 per cent. Gross differences, however, varied from about 4 per cent up to 9.5 per cent. Differences in total land in farms were of approximately the same size. Net difference varied from 0.6 per cent up to 4.8 per cent and gross differences from 5.4 per cent up to 10.9 per cent.

The data indicate that 1.3 per cent of the population was missed in the original enumeration. The percentage of persons missed varies from a low of 0.2 per cent to a high of 2.3 per cent. This is an improvement over an estimated 3.3 per cent missed in an earlier census pretest done in the spring of 1948. The difference probably reflects special efforts to improve quality in the October pretest.

The differences reported occurred in spite of heavy emphasis upon coverage in training the enumerators. It is obvious that only the most strenuous efforts can prevent such coverage errors.

TEST NARRATIVES

Another testing device that has been used in preparatory work for the 1950 Censuses is the "test narrative." An interview is read to

enumerators who record it on regular schedules. This is a relatively cheap method of (a) investigation of some of the advantages and disadvantages of schedules having different formats, and (b) comparing the performance of enumerators in a controlled situation with their performance in a regular census enumeration.

Use of these narratives has been encouraging in that the errors obtained under such a procedure give us some insight into how schedules should be designed in order to reduce recording errors, and also point up areas where special training is needed. Furthermore, although scoring test narratives presents considerable difficulties and is somewhat arbitrary, such "scores" may predict quality of performance in the census (performance here being measured by the proportion of discrepancies discovered for each enumerator in the quality check). In one of the enumerations in which the test narratives were used, enumerators scoring in the highest and those scoring in the lowest quartile were compared: On about half the items the differences were extremely small and in some the low scoring enumerators were even slightly superior. But in several items—number of rooms, monthly rent, date of birth, educational attainment, weeks worked the previous year and wages or salaries—the high scoring enumerators had 5 or more percentage points less discrepancy than the low scoring enumerators and there were no items in which differences of 5 or more percentage points occurred in favor of the low scoring enumerators.

IMPLICATIONS AND PROBLEMS

1) The results thus far obtained, while far from conclusive, point to the importance of attaching measurements of non-sampling or response error to the more important Census figures. The Bureau has, insofar as its resources permitted, attempted to indicate in releases and publications the major limitations on the figures published. Naturally, we have not been able to describe every condition affecting the accuracy of each of the millions of figures published. It is important that users of Census data realize that getting usable data on migration, on fertility, on farm expenditures, etc., involves more than putting a question on the Census schedule and grinding out figures. A great deal of thought goes into the advisability of including in the Census each of the characteristics proposed. In those considerations increasing weight is being placed upon the feasibility of accurate measurement in determining whether an item should be included. For many purposes, it can be argued that some figure is better than no figure. On the other hand, there is obviously an error level at which a guess is as good as the figure

which can be gotten. There is, of course, no ironclad rule for determining when a figure is too inaccurate to be useful. An error of 4 per cent in the population count of the United States might impair the usefulness of the figure for many important purposes. An error of 100 per cent in the figures on oats produced in Hillsboro County, Florida, might be perfectly satisfactory if our main interest is in being sure that oat production in the county is small. An error is "big" or "small" only in relation to the uses of the data. This fact should be kept constantly in mind in planning a census or any other survey.

There is also an extremely important implication of response error measurement for sample design. In many cases, sampling ratios are set at the level necessary to insure a small sampling error. If planners were to investigate their response errors, they would find that, in many cases, efforts devoted to the attainment of a small sampling error are in large part wasted because of failure to reduce response errors substantially larger than the sampling error. It is, for example, somewhat ridiculous to spend $10,000 to reduce sampling error from 3 per cent to 1 per cent when the net response error of a complete count is 10 per cent.

2) The pitfalls of *a priori* assumptions about a survey method have been suggested above. We cannot overstress the need for empirical test of assumptions. This does not mean that we would advocate purely empirical methods. If improvements in techniques are to be made, we must combine a sound theoretical formulation with rigorous empirical tests.

3) The use of a quality check technique supplies very useful information about the accuracy of census figures. It must be realized, however, that the method has definite limitations. Other methods for measuring response error have been used and can be explored more fully. One technique is the checking of reports against outside criteria. For example, occupation and industry might be checked against employer's records, education against school records, and cotton production against cotton gin records. Checks of this type may be expensive when independent check data are not readily available, although they are not prohibitively expensive.

It should also be noted that these checks can be very disappointing. In a recent (extremely small scale) attempt to check census pretest reports of age and birthplace against birth certificates there was a substantial proportion of the cases where no birth certificate could be located. One possible reason for failure to find a birth certificate is that

none exists, but another possibility is that data used for the search were inaccurate. If, for example, age as reported in the census was in error by more than 5 years, the method of search would make it unlikely that a birth certificate would be found. Similar problems arise in other types of record checking and may well invalidate the results completely.

There are characteristics where checks against outside criteria are nearly impossible because of the lack of satisfactory outside criteria. This is true of items such as residence a year ago and value of farm land and buildings. Going outside the case of census surveys, checks against outside criteria, for example, are impossible for many opinion items.

4) Even where satisfactory techniques for measuring error exist, the problem of experimental design is not simple. Attempts to control *all* the factors involved seem to be futile. Perhaps the only satisfactory method of design is a thorough and controlled randomization. Furthermore, the large differences between counties and between enumerators indicate that clustering of cases and use of relatively few enumerators will not give efficient discrimination among alternative procedures. Future experimental design should be aimed at more complete randomization with a fairly large number of small areas and many enumerators. Resources may not permit this type of design under conditions such as those of the October 1948 census pretest where it was necessary to estimate costs, and completion rates, and to take into account administrative problems and the effects of publicity.

5) The measurement of response errors is, of course, only one phase of a program of response research. As a measure of accuracy of data, the "global" response error is satisfactory and is, in fact, the most important measure. A program aimed at control and reduction of error must, however, adopt more analytical methods. Sources of error must be isolated and measured separately. It is necessary to determine how much error is attributable to the respondent and how much to the interviewer and within these categories what are the effects of question wording, schedule design, definitions of terms, enumerator training, etc. Even on some of the most elementary points, knowledge is inadequate. Some features of schedule design which, on the basis of past experience in other fields, would be considered unsatisfactory seem to make very little difference in the results obtained by the Census Bureau. The field for exploration is large and the work in it, while impressive from some standpoints, is still in its infancy.

In conclusion, we wish to note that the material discussed in this

paper represents the work of many persons and groups in the Bureau of the Census. The planning and development of even the smallest of the surveys mentioned is an intricate job involving the intimate cooperation of several divisions of the Bureau and of outside agencies.

REFERENCES

[1] Deming, W. Edwards, "On Errors in Surveys," *American Sociological Rievew*, pp. 359–369, Vol. 9, No. 4, August 1944.

[2] Ducoff, Louis J. and Bancroft, Gertrude, "Experiment in the Measurement of Unpaid Family Labor in Agriculture," *Journal of the American Statistical Association*, 40: 205–213 (June 1945).

[3] Bancroft, Gertrude and Welch, Emmet H. of the Bureau of the Census, "Recent Experience with Problems of Labor Force Measurement," *Journal of the American Statistical Association*, 41: 303–312 (September 1946).

A NOTE ON THE DERIVATION OF INCOME ESTIMATES BY SOURCE OF INCOME OF PERSONS MAKING LESS THAN $500 PER ANNUM, 1944–1948*

ULRIC H. WEIL†

This paper derives estimates of money income received by persons making less than $500 per annum for the years 1944–1948. The estimates are classified by source of income and are obtained by adjusting basic census data for various types of understatement.

INTRODUCTION

PROCEDURES to estimate money income received by persons by size of income are still in an embryonic stage, although it is unquestioned that adequate income size distributions are basic to the study of many important economic problems. The reason for this is not the lack of determination and skill on behalf of the investigator but rather the paucity or non-existence of adequate data. In this paper, an attempt has been made to derive estimates by source of income of money income received by persons making less than $500 for the years 1944 to 1948. Because of the above mentioned lack of adequate data a rather intricate method of adjusting basic census estimates[1] was created and this method is fully described in the following pages. The adjustments made broadly speaking allow for understatement of frequencies, understatement of income reported in the respective size class and shifting of frequencies between size classes on account of 'the income blow-up.'[2] It may be submitted here that this or a similar approach could be utilized to build up a complete income size distribution which, though far from being beyond reproach, may be considered to represent an improved method of deriving size distributions of income received by persons. The limitations of some of the existing size distributions partly because of the restrictions inherent in the data from which they were derived are too well-known to require examination in this paper.[3]

The motivation for choosing the lowest size class as the subject matter of this paper lies in the desire to adjust Department of Commerce

* Acknowledgment should be expressed to Mrs. S. Goldsmith and Mr. M. Liebenberg of the Office of Business Economics, Department of Commerce, for their generous cooperation and criticisms extended in the preparation of this paper.

† Fellow of the Social Science Research Council, 1949–1950.

[1] Bureau of Census Publications: P-5 No. 32; P-60 No. 3, 5; P-50 No. 15.

[2] This term refers to the procedure used in inflating results obtained via sampling methods by a proportionality factor. This factor relates the sample size to the size of the universe.

[3] See "An Appraisal of Field Surveys of Consumer Income" by R. Wasson, A. Hurwitz, and I. Schweiger, CRIW, April 1949.

consumer money income estimates among others for the income of persons making less than $500. This adjustment appears essential when comparing Commerce data with similarly conceptually adjusted Bureau of Internal Revenue data, in order to determine the 'gaps' between these two series. Idealistically speaking, i.e., among others in the absence of under-stated or non-stated reportable income—such a 'gap' should be non-existent, or at least be very small, because in the years under consideration here the filing requirement was in terms of gross income of $500 or more. The recent congressional emphasis upon the economic welfare of the "lowest third" of the income-earning population also serves to stress the need for more empirical information concerning the frequency, type of income, and its total amount which accrues to those persons finding themselves in the lowest brackets of the income size distribution.

TABLE I

ADJUSTED MONEY INCOME BY SOURCE FOR STATED YEARS OF PERSONS 14 YEARS AND OVER MAKING LESS THAN $500 BY SOURCE OF INCOME

(Figures in Billions)

	1944	1945	1946	1947	1948
Total Income	$4.40	$3.90	$3.86	$3.78	$2.84
Civilian Wages and Salaries	3.31*	2.69*	2.83*	2.81*	2.14*
Non-Farm Entrepreneurial Income	.11*	.16*	.23*	.17*	.13*
Farm Entrepreneurial Income	.12*	.13*	.09*	.13*	.10*
Military Pay	.25†	.47†	.24†	.22†	.11†
Rents¶	.29§‖	.24‡‖	.26§‖	.25§‖	.20§‖
Interest and Dividends	.09†	.05†	.06†	.07†	.05†
Social Security Payments	.14§	.12†	.12§	.11§	.09§
Other Receipts	.09†	.04†	.04†	.02†	.02†

* Estimates obtained by multiplying field survey money average times number of earners as per census adjusted to include missing frequencies (or proportion thereof) for the size class $1–499 as per table 13 of Mrs. Goldsmith's paper.

† Obtained by using a percentage derived from table 4 (source item to Wages and Salaries) in Mrs. S. Goldsmith's paper.

‡ Obtained from BOC field survey unadjusted.

§ Obtained from BOC field survey unadjusted by using same proportion to total salaries and wages as observed in 1945: (.0895 for rents; .0423 for social security payments) for all years.

‖ Rental income is composed of rents and income from roomers and borders on the basis of census data for 1945 for the size class $1–499 the proportions in which these categories are contained in rental income above are .52 and .48. These ratios have been used for all year.

¶ Including roomer-boarder income.

Rent Excluding Roomer and Boarder Income

(Figures in Billions)

1944	*1945*	*1946*	*1947*	*1948*
.15	.125	.135	.13	.105

Sources: Bureau of Census Publications; P-5, No. 32, P-60, Nos. 3, 5, P-50, No. 15; Goldsmith, Selma F. *Appraisal of Basic Data Available for Constructing Income Size Distributions.*

Census income estimates derived from field surveys for the years 1944–1948 are basic to the data shown in Table I. A detailed description of the sampling technique and design employed in these surveys may be found in the Census publications already referred to above. However, since the estimates sought here will eventually be applied to Commerce data, in order to make those data comparable with Internal Revenue data after other appropriate conceptual adjustments have been completed, a brief list of the differences between the Census and Commerce procedure of income estimation is given below:[4]

1. The definitions of income of the two agencies differ. Commerce includes certain items which are not included in the Census definition, the most important of which are income in kind and net change of farm inventories. Census includes certain items which are not included in the Commerce series, chiefly contributions for support received from persons who are not members of the household. It is assumed by the writer that for the $1–499 size-class the conceptual differences after cancellation of compensating differences result in a net difference in favor of Commerce concepts, i.e., Census estimates in general and especially for this size class are too low.

2. Commerce includes the income of all persons in the continental U.S.A. during any given year, while Census excludes from its sample income of persons in institutions and quasi households. Again this may tend to make Census estimates too low in terms of the Commerce concept.

3. Commerce data are estimated largely on the basis of such record data as Old Age and Survivors' Insurance wage records, statistics of income, governmental payments, etc. while Census data are derived from field surveys. It is generally felt that, because of the reluctance of respondents to give income information and/or their propensity to forget certain irregular receipts of income, field survey income estimates tend to underestimate income aggregates.

4. In addition to the above the writer feels that on the basis of more recent investigations,[5] it may well be argued that the independent control factor (persons 14 years of age and over in the U.S.A.), to which the sampling frequencies are 'blown-up' by Census, has been consistently and considerably under-estimated.

All in all, then, the raw Census estimates may be assumed to be too low on the basis of significant conceptual differences between Census

[4] For a more detailed statement see the BOC publication: P-S No. 22.

[5] See the paper by Mrs. S. Goldsmith of the Department of Commerce, "Appraisal of Basic Data for Constructing Income Size Distribution," CIRW, April, 1949, especially tables 11–13.

and Commerce series, especially points (4), (3), and (2) above. Furthermore, when attempting to utilize Census income estimates for the years 1944–1948 it should be remembered that only the 1945 figures represent actual money incomes broken down by source of income as ascertained in a field survey. Total money income figures for the other years can be obtained only by multiplying the mean income of the $1–499 size class by the frequency of that size class for the given year. A breakdown by source of income in those years (i.e., 1944, 1946–1948) had to be achieved by applying either the proportions indicated by the 1945 Census source pattern or by the source pattern given by the Department of Commerce in the other years (see footnotes to Table I). To make allowance, then, for the undercoverage of the Census universe and the understatement of the income given by Census an attempt could and should, therefore, be made here to adjust census data (1) for the undercoverage of the universe (persons 14 years of age and over having incomes of $1–499)—points (1), (2), and (4) above—and (2) for underreporting of income by members of the universe because of the "memory factor," i.e. their unwillingness to state their full income or their refusal to divulge any—point (3) above. The first adjustment which involves merely a "stepping up" of the frequencies per size class as reported by Census, has been carried out here and its methodology is described below. The second adjustment, which requires an allowance for shifting of members within size classes themselves, involves the task of augmenting an income size distribution by the relative amount of the 'missing' income. Its methodology is also outlined in the following pages.

I. THE ADJUSTMENT FOR STEPPING UP THE FREQUENCIES IN THE $1–499 SIZE CLASS

The following reasoning underlies this adjustment: in Table 13 of Mrs. Goldsmith's paper revised and adjusted estimates of the number of civilian earners 14 years old and over for 1944 to 1946 by source of earnings are built up from Old Age and Survivors' Insurance data and other sources of data (for a detailed statement of the procedures followed by Mrs. Goldsmith in deriving these estimates see the notes to Tables 10–11 of her paper). By comparing these estimates with the corresponding Census estimates 'missing' frequencies for the aggregate income distribution were obtained subclassified by earners with wages and salaries (with farm or non-farm income) and earners with no wages and salaries but with farm or non-farm income. For 1947 and 1948 the same procedure was followed using Census estimates and Commerce

data as far as available to which 1946 relationships were applied to obtain the necessary sub-classifications by type of income recipient. It is clear that by not including in the computation of persons with incomes of $1–499 those persons deriving income exclusively from such sources as Rents, Interest and Dividends, Military Payments, Social Security Payments and Miscellaneous Receipts the total frequency of 'income recipients' for this size class is understated. However no information seems to be available to bridge the gap between earners and income recipients frequencies. It would appear, however, that the number of people having exclusive income from non-earnings is very small in the size class under consideration. By attributing to the 'stepped-up' earner frequencies unearned income in given proportions as derived from Census relationships in 1945 and proportions derived from Table 4 in Mrs. S. Goldsmith's paper for the other years, in order to obtain total money income, the shortcoming of the above described procedure of jumping from earners to income recipients without adding on actual frequencies of persons having income from exclusively unearned sources seems to be minimized.

Table II, then, reproduces the information necessary to undertake the first of the two adjustments required to obtain money income aggregates by source for the size class $1–499 for 1944–1948. In order to convert the frequencies shown in column 9 of Table II into money figures, they were multiplied by the class income source average as derived from Census data for 1945. Since only in 1945, reliable money income and income recipient frequency information by source is available, these 1945 averages were also used for all other years.

	1945 Averages
Salaries and wages	$226.7
Farm earnings	156.
Non-farm earnings	189.

On the basis of the relationship of interest and dividends, other payments and military payments respectively to civilian wages and salaries as derived from Table 4 of Mrs. S. Goldsmith's paper for the years 1944–1948, income aggregates were added representing income from these sources received by the frequencies belonging in the $1–499 size class. A similar procedure was followed for rents and social security payments except that for these income sources the 1945 Census relationship between the specified source and the civilian wages and

TABLE II

ADJUSTMENT FOR STEPPING UP CENSUS FREQUENCIES OF INCOME REPRINTS IN THE $1–499 SIZE CLASS

Year	*Census proportions of income recipients $1–499 size class*	*Census frequencies of income recipients $1–499 size class* (in millions)	*Proportion of all specified earners to total civilian earners as per Selma Goldsmith's paper—table 11*			*Missing frequencies as per table 15 in Mrs. Goldsmith's paper* (in millions)			*Total frequencies $1–499 size class (in millions)* (2)+(6)+½[(7)+(8)]
			Wages and Salaries	Farm Entrepreneurial Income	Non-Farm Entrepreneurial Income	Wages and Salaries‡	Farm Entrepreneurial Income§	Non-Farm Entrepreneurial Income§	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1944	.12	12.18	.872	.065	.059	11.5	.8	.4	24.3
1945	.13	10.28	.868	.065	.064	11.1	1.3	1.6	22.8
1946	.18	11.85	.864	.062	.073	8.7	—	1.0	21.1
1947	.18	11.0	.864*	.065*	.064*	8.5*	1.0	1.22*	20.5
1948	.17*	11.2†	.864*	.065*	.064*	4.25*†	.5*†	.61*†	15.9

* 1945 relationships were used, because of lack of information.

† A basic estimate of total earners of 70 millions was derived from the Survey of Current Business. This figure was used to obtain the missing frequencies between Census and Commerce.

‡ Column (6) indicates the share of total 'missing' frequencies between Census and Commerce earner estimates (table 13—Mrs. S. Goldsmith's paper) allocated to civilian wage and salary earners in the $1–499 size class. The assumption was made that all 'missing' civilian salaries and wages frequencies belong in this class.

§ Column (7, 8) give similar data of 'missing,' 'farm,' and 'non-farm' earners. However here only 50% of the 'missing' frequencies was assumed to belong to the $1–499 size class.

salaries was utilized. The above procedure, then permits the estimation of money income by source of 'income recipients' making less than $500.

However, after completion of this first adjustment, the 'gap' between Commerce total income estimates and frequency adjusted Census income estimates is not yet filled. The following figures indicate the proportion which 'adjusted' Census estimates now represent relative to Commerce estimates.[6]

Year	Proportion of Adjusted Census Estimates Relative to Commerce Estimates
1944	.808†
1945	.77†
1946	.802†
1947	.82†
1948	.81*

* Estimated.
† See Table 2 of Mrs. S. Goldsmith's paper.

II. AUGMENTING AN INCOME SIZE DISTRIBUTION BY STIPULATING A CONSTANT LORENZ CURVE

It appears from the above proportions that the 'adjusted' Census estimated have to be further adjusted, because even after adjusting for missing frequencies in the $1–499 size class the income aggregates as derived by Census and Commerce are far apart. It is assumed that this proportion of income missing from 'adjusted' Census estimates relative to Commerce estimates is accounted for by the 'memory factor' mentioned in point (3) above. This 'forgotten' income, therefore, must be added to the 'adjusted' Census estimates. This has been done for the $1–499 size class on the assumption of a constant Lorenz curve. The shortcomings and underlying assumptions of these methods are fully discussed by Messrs. Liebenberg and Kaitz.[7] The second adjustment therefore has been carried through utilizing the constant Lorenz curve assumption in the face of these shortcomings in view of the lack of adequate supplementary data and other basic information, which might throw adequate light on the pattern of under-reporting of income by source and size of income, on the basis of which more satisfactory adjustment methods could be applied.

[6] This statement and the supporting data, of course, imply the validity of the assumption of equal percentage undercoverage over the whole range of the income distribution.

[7] See the paper of Messrs. M. Liebenberg and H. Kaitz "Computing an Income Size Distribution from Income Tax Data, 1944," CRIW, April 1949, pp. 1–4.

In accordance with the above, therefore, the total money income for the size class $1–499 was redistributed along a frequency density function, the class limits of which were interpolated for on the basis of the formulae on page 83 of Messrs. Liebenberg and Kaitz's paper.[8]

The final income estimates for the size class $1–499 for the years 1944–1948 adjusted as described above are given in Table I. Their absolute magnitude—aside from their further important use in connection with Bureau of Internal Revenue and Commerce income series comparisons—plus the associated frequencies shown in column (9) of Table II may give some empirical basis for determining the relative importance or unimportance of this lowest income size class.

[8] The formula given there reads: "$a = 2FKZ^2(1Z) - Z^2(3-4Z)(A - FX_0) + fX_0$ where f = the frequencies in the interpolated portion of the interval; a = the aggregate income in that portion; F = the frequencies in the entire interval and A the corresponding aggregate; X_0 = the lower limit of the class; K = the size of the interval; and $Z = X - X_0/K$, and X is the income at the point of interpolation."

ESTIMATING PRECISION OF MEASURING INSTRUMENTS

H. Fairfield Smith
Rubber Research Institute of Malaya

This paper presents a technique for computing the precision of two measuring instruments when there is a linear relation between the scales of the two instruments.

Grubbs (1948) has considered the problem of estimating precision of measuring instruments under the following conditions: The test is destructive so that on any given piece of material only one test can be made, but it can be observed simultaneously by several instruments. The material is variable so that one cannot obtain replicate observations with the same instrument which will show directly the random variance of the instrument readings. The problem is to obtain separate estimates of the variances of the instrument error and of variability of the material under test.

The procedure elaborated by Grubbs implies the assumption that any systematic difference between the instruments is constant throughout the range of observation. This restriction has however not been explicitly stated and it appears advisable that it should be pointed out since it may in some circumstances be important.

In the example given by Grubbs, the range of observations is so small that it is unlikely that the restriction could there introduce appreciable error. But with a greater range of material this might not be so; it is quite reasonable to suppose that one operator's delay in stopping his clock might increase or decrease relative to that of another operator as the total recorded time is altered. (In this example the "instruments" were in fact abilities of persons to operate the instruments. The data show that the major alteration of systematic errors throughout these experiments was due to improvement with practice, and the alterations thus introduced would mask any association with magnitude of the observation.)

More generally, consider that the problem is to compare two instruments for measuring a character which can only be observed *via* these instruments, so that there are no theoretically absolute values against which the instruments can be calibrated; the definition of the character being in fact determined by the preferred method of measurement. (For example: suppose no quantitative measure of temperature had yet been invented and we had just conceived the idea of basing one on

alteration in volume of one or other of several liquids; one would have to consider the possibility that the thermal expansion of two different liquids might not always be proportional and the scale might differ according to which one should be adopted as standard.) In these circumstances when comparing two (or more) proposed instruments it would be necessary to consider both accuracy and scale jointly.

Let observed readings on one instrument be designated $x=\xi+\epsilon$ where ξ is the true value (on the scale of this instrument) for the variate observed and ϵ is a random error of observation. Similarly denote the readings on another instrument by $y=\psi+\eta$. For simplicity let us assume that the scales are linearly related by

$$\psi = \alpha + \beta\xi$$

(Extension of the argument to a curvilinear relationship will be obvious.) Following the same argument as in Grubbs' paper we obtain from simultaneous observations on several test pieces

$$\begin{aligned} V(x) &= V(\xi) + V(\epsilon) = \beta^{-2}V(\psi) + V(\epsilon) \\ V(y) &= \beta^2 V(\xi) + V(\eta) = V(\psi) + V(\eta) \\ \text{Cov}\,(xy) &= \beta V(\xi) = \beta^{-1}V(\psi) \end{aligned} \qquad (1)$$

To obtain solutions for $V(\epsilon)$ and $V(\eta)$ from these equations the most practical, and probably the most efficient, method appears to be to estimate β by the method given by Wald (1940).[1]

Example: A new instrument for observation of shear stresses having been developed, it was required to ascertain whether or not it gave sufficiently greater accuracy to justify the extra labour required to use it instead of the older and simpler instrument. A number of paired observations were made, 31 pairs on samples of a timber (A) with high shear stress, 13 pairs on samples of a timber (B) with low shear stress, with the following results, where x stands for observations with the new, y for the old instrument:

---

[1] I have tried one or two other procedures. For example, by dividing the data into two sections, with variances V and V', etc., and assuming that the ratio $V(\epsilon)/V(\eta)$ is the same in both, we get

$$\frac{V(x)-Cov(xy)/\beta}{V(y)-\beta Cov(xy)} = \frac{V'(x)-Cov'(xy)/\beta}{V'(y)-\beta Cov'(xy)}$$

which, when solved for β, gives an equation similar in form to Kummell's equation for the "line of best fit" (quoted by Edwards Deming, 1944); but such procedures appear to give inefficient estimates of β. Kummell's formula itself cannot of course be used to estimate β since it requires *a priori* knowledge of the ratio $V(\epsilon)/V(\eta)$ which we want to estimate.

Series	Mean difference $(\bar{x}-\bar{y})$	Standard deviation of $(x-y)$	Mean
A	−44	195.6	2021
B	110	72.7	940
Difference	154 ± 40.5		

The difference in systematic error of the two instruments in the regions covered by the two series is highly significant. Furthermore errors of observation were found to increase approximately in proportion to the stress. For further analysis it was therefore deemed advisable to stabilize variances by working with logarithms of observations; and for simplicity, in the absence of more complete information, it was postulated that the scales of the two instruments might be related by

$$\log \psi = \alpha + \beta \log \xi$$

To estimate β by the "method of averages" it appears most logical to consider the means of each series, since this fulfils the condition that the grouping should be independent of the errors ϵ, η, as well as maximizing the difference of the means. The means are

Series	$\log \bar{y}$	$\log \bar{x}$
A	3.30807	3.29943
B	2.94313	2.99615

whence, following Wald with necessary modifications for unequal numbers of observations in the two groups, we estimate

$$\log y = -.6618 + 1.2032 \log x \pm .0473$$

Then, inserting this estimate of β in equations (1), along with variances and covariances within groups,[2] we estimate $V(\eta)$ and $V(\epsilon)$ as follows:

Series	A	B	Pooled
d.f.	30	12	42
$V(x)\times 10^3$	1.226	1.704	1.363
$V(y)\times 10^3$	2.045	3.728	2.526
$\mathrm{Cov}(xy)\times 10^3$	.760	1.900	1.086
$V(\epsilon)\times 10^3$	.593	.121	.460
$\beta^{-2}V(\eta)\times 10^3$	.787	1.005	.842

We compare $\beta^2 V(\epsilon)$ to $V(\eta)$, since obviously the errors are only of interest relative to distance between corresponding points on the two scales.

Note on the example given: Actually the work from which the above data have been taken differs materially from the situation considered by Grubbs, in that the two instruments cannot be put on to breaking the self-same piece of timber. The data quoted are for observations on matched pairs of test pieces. $V(\eta)$ and $V(\epsilon)$ therefore include variation between matched pairs as well as instrument errors. These two sources of error cannot be separated, but by using matched pairs variability of material has been made as small as will be possible in any practical problem. Failure to separate these sources of error is immaterial in judging the values of existing instruments to discriminate between samples of timbers. It would matter if the problem were to decide whether or not it might be worth while to devote considerable effort to producing yet another instrument with still lower intrinsic error.

In such problems attention to scale is important as it may frequently be modified by operational procedure of each instrument. For example: in testing tensile strength of rubber, both shape of test piece and rate of separation of the grips may modify the stress-strain curve. There seems no reason to suppose that analogous considerations would never apply in the problem considered by Grubbs, although none such are known to me.

In all similar problems with which I am acquainted separate test pieces are necessary for each observation. In this situation clearly the instrument error plus minimum test piece variability can be obtained by running a series of observations with the same instrument on matched pairs, and the use to be made of paired observations with alternative instruments can be limited to comparing their scales. The analogy to Grubbs' problem was raised by the condition that when this work was first brought to my attention the data reported above had been obtained, without comparable series on the same instrument, and an opinion was required before further work was done.

As far as could be judged from the data originally presented the hypothesis regarding variable scale as set up above seemed reasonable. (One reason being that assumption of equal scales leads to a negative estimate of $V(\epsilon)$ in series B.) Actually later work has shown that both

[1] In getting the average values equations (1) are applicable either to pooled variances within groups or to total variances over all observations. The resulting estimates of $V(\epsilon)$ etc. will differ in the ratio of the degrees of freedom associated with the estimates of variance. Wald shows that the d.f. available for estimating $V(\epsilon)$ and $V(\eta)$ are n_1+n_2-2; it follows that the within groups variances are appropriate for obtaining the pooled estimates.

instruments appear to have equal scales, but may give divergent results on particular species of timber at any point on the scale. The above data are therefore to be taken only as illustrative of a situation which may arise elsewhere. They do not provide the correct solution in this particular case.

ACKNOWLEDGEMENT

I am indebted to Mr. A. V. Thomas of the Malayan Forestry Service for permission to use his data.*

LITERATURE CITED

Deming, W. Edwards (1938, 1944). *Statistical adjustment of data.* J. Wiley & Sons, New York.

Grubbs, F. E. (1948). On estimating precision of measuring instruments and product variability. *Journal of the American Statistical Association* 43: 243–264.

Wald, A. (1940). The fitting of straight lines if both variables are subject to error. *Annals of Mathematical Statistics* 11: 284–300.

---

* EDITORIAL NOTE: The following references, which were not available to the author, seem relevant:

Bartlett, M. S. (1949). Fitting a straight line when both variables are subject to error. *Biometrics*, 5, 207–212.

Toner, R. K., Boner, C. F. and Whitwell, J. C. (1948). Moisture determination in textiles by electrical meters. Part I. *Textile Research Journal*, 18, 526–535.

Bartlett's paper gives a modification of Wald's method for estimating the slope β, while the paper of Toner et al., contains the basic equations (1) used in the present paper.

HORATIO MILO POLLOCK, 1868–1950

Dr. Horatio Milo Pollock, a fellow of the American Statistical Association and a vice-president in 1933, died in Petersburg, Virginia, on May 8, 1950 at the age of 81. Dr. Pollock had retired from his position as Director of Statistics in the New York State Department of Mental Hygiene in 1944, after more than three decades of distinguished service.

He was born in Patria, Schoharie County, New York on September 2, 1868. He labored on his father's farm throughout his early life, as a result of which his formal schooling was delayed. Once started, however, he advanced rapidly. He completed his high school courses in record time. He received the degree of Bachelor of Science from Union College in 1895. This was followed by a period of post-graduate study in biology in Germany, where he received the degree of Doctor of Philosophy from the University of Leipzig in 1897. During the same year he received the degree of Master of Science from Union College. Half a century later, he received the honorary degree of LL.D. from his alma mater.

For fourteen years following his return from Germany, Dr. Pollock divided his time between teaching and civic activities. At various times he taught science, sociology and economics at the New York State College for Teachers in Albany and at Union College. For four years he was secretary of the Civic League of Albany. For seven years he was senior examiner for the New York State Civil Service Commission.

In 1911 the New York State Commission in Lunacy sought a statistician, and the choice fell upon Dr. Pollock. At that time the statistics of mental diseases were in a chaotic state. Several distinguished physicians, not primarily statisticians, had written scattered papers from time to time, but nobody had cultivated the field assiduously. When the State Commission in Lunacy was organized in 1889, it recognized an imperative need for complete and accurate statistics. The Commission hoped that such statistics would be the basis for administrative and scientific policies. But the program did not get under way until 1911, when, with Dr. Pollock's appointment, new life was injected, and the program quickly assumed shape.

Dr. Pollock was successful in establishing the first file dealing completely with hospitalized patients with mental disease. Within a few years the file had grown to such proportions that it was possible to begin a continuous flow of studies dealing with the number and char-

acteristics of such patients, their annual accretions and the outcome of treatment.

His achievements brought him to the attention of the country at large. When the National Committee for Mental Hygiene undertook annual surveys of patients in mental hospitals, Dr. Pollock became consultant to the Committee. During the first World War, he served as consulting statistician to the office of the Surgeon-General. Subsequently he became a consultant to the United States Veterans Bureau. In 1920 he helped the State of Illinois set up a statistical system. From 1921 to 1926 he was a special adviser to the Census Bureau with respect to statistics of mental disease, mental defects and epilepsy. For many years he was chairman of the Committee on Institutional Statistics of the American Statistical Association.

He was the author of *Mental Disease in the United States as Shown by the Federal Census of 1923*. This marked a transition from a decennial to an annual census of such patients. The report differed from its predecessors in two important respects. It introduced a classification of mental diseases which had been adopted by the American Psychiatric Association, and it differentiated first admissions from readmissions. These remain the basic classifications. Subsequently Dr. Pollock was instrumental in the preparation of similar classes of statistics for institutions for mental defectives and epileptics.

These statistical activities represented only part of Dr. Pollock's contributions. For many years he edited the publications of the New York State Department of Mental Hygiene, including the *Psychiatric Quarterly* and the monthly publication entitled *Mental Hygiene News*. He was a leader in the introduction of occupational therapy, and the use of family care for patients with mental disease. He was editor and co-author of the first detailed publication in the United States dealing with the principles of family care. His interest in the care of mental defectives was recognized by election to the presidency of the American Association on Mental Deficiency in 1942.

Dr. Pollock's last years in the Department of Mental Hygiene were devoted in large part to the planning and editing of a monumental history in four volumes on the care of the mentally ill in New York State from colonial days to the present, including a detailed description of the institutions within the Department of Mental Hygiene. This history attests to his devoted services to the mentally ill for almost four decades, and when published will honor his memory, and keep it fresh in the minds of his many friends and associates, within and without the New York State Department of Mental Hygiene.

BENJAMIN MALZBERG

BOOK REVIEWS

(Pending the establishment of a new editorial board, the review section has been edited for the Secretary's Office since August 1949 by Dr. Ernest Rubin.)

Experimental Designs. *William G. Cochran* (Professor of Biostatistics, The Johns Hopkins University, Baltimore, Md.) and *Gertrude M. Cox* (Director, Institute of Statistics, University of North Carolina, Raleigh, N. C.). New York 16: John Wiley and Sons, Inc. (440 Fourth Ave.), 1950. Pp. ix, 454. $5.75.

REVIEW BY PALMER O. JOHNSON
University of Minnesota

THE authors have written this book primarily for the experimenter, who will find here the description of the most useful designs that have been developed up to this time. It is a handbook to which he may refer for intelligent guidance when he contemplates an experiment. The experimenter should possess, besides first-hand knowledge of the problems in his field, a working knowledge of the analysis of variance. With this background he opens a book unexcelled in its exposition of the scientific and statistical basis of the principles of experimentation and their transition into designs. On the maxim that example is better than precept, an unusually successful attempt has been made to exemplify the processes by a comprehensive selection of examples. These are for the most part drawn from biology, particularly agriculture, in which field the greatest part of the experience of the authors has been gained.

Chapters 1, 2, and 3 concentrate on topics of general importance, including the mathematical formulation of the assumptions and the principles of statistical inference leading from the assumptions to the statistical method. Chapter 1 treats of the contribution of statistics to experimentation, particularly to the problem of induction as developed by the theory of statistics. The problems of testing hypotheses and of estimation are thus introduced and elaborated. The fundamental role of randomization is stressed in making unbiased estimates of error possible and in neutralizing the effects of factors whose incidence is independent of the treatments. Chapter 2 is devoted to methods for increasing the accuracy of experiments: (1) increase in the size of the experiment either by more replicates or treatments, (2) refinement of experimental technique, and (3) handling of experimental material so as to reduce the effects of variability. Useful tables, based on the statistical solution of the problems, are made available, such as the table for determining the number of replications needed for a given probability of obtaining a significant result.

Chapter 3, "Notes on the Statistical Analyses of the Results," is perhaps the most fundamental and informative chapter. The general technique which governs the analysis of all designs, that is, the method of least squares, is described and applied to one of the simplest designs, that of randomized

blocks. Some account is also given to such special statistical techniques as the analysis of covariance and of experimental results from which the data from one or more experimental units are missing.

Beginning by writing down the mathematical model, the assumptions made in the model are then succinctly stated. The best estimates of the unknown parameters in the model are then determined. The role of the analysis of variance is carefully developed and the complete analyses discussed and illustrated. Particularly important is the clear explanation of the subdivision of the sum of squares for treatments, of the calculation of the standard error for comparison among treatments, and of the subdivision of the sum of squares for error. Finally, the effects of errors in the assumptions underlying analysis of variance are discussed with some attention to the use of transformations and techniques of analysis that may be useful in dealing with the situations encountered in practice. The statistician would likely desire a more comprehensive discussion of the analysis of variance including its use in estimating the several variance components. Some consideration might well have been given in this chapter to the role of presupposition in the design of experiments.

If the content of the first three chapters has been mastered, the reader is well prepared for the sequel. Chapters 4–13 deal, respectively, with completely randomized block and latin square designs, factorial experiments, confounding, split-plot designs, factorial experiments compounded in quasi-latin squares, balanced and partially balanced incomplete block designs, latin and cubic lattice designs, balanced incomplete blocks, lattice squares, and incomplete latin squares. The procedure is to present the computations with a worked numerical example for each of the designs likely to be used most frequently. There are accompanying plans for each major design and a specification of the experimental situations for which each design is most suitable. All told, there are plans for 150 of the designs found most useful.

Chapter 14 deals with the fundamental problem of the analyses of the results of a series of experiments. The final chapter gives useful tables of the random permutations of 9 and 16 numbers.

This book will set the standard for a long time to come. Mathematicians have, at least since the time of Euler, found much interest in combinatorial problems dealing with the arrangement of things in configurations of various kinds. It remained for R. A. Fisher to show how to put these combinatorial principles to work. The principles of experimentation have stood the test of time and are finding increasing application in many fields of science. This book will undoubtedly serve to increase and elevate the standards of their use and in doing so demonstrate the most important practical application so far made of statistics.

Statistics without Numbers, a Visual Explanation. *F. Kafka* (The City College, New York). New York: Lifetime Editions, Incorporated, 1950. Pp. x, 111. $2.00.

REVIEW BY WILLIAM C. JAMES

Director, Statistical Division, National Safety Council

THIS book undertakes to interpret the basic ideas of the science of statistics by a combination of simple, straightforward text and roughly sketched but nevertheless clear pictographs. The underlying principle is oversimplification. Its appeal is to the unsophisticated consumer of statistics, and this fact is expressly stated in the preface. Any evaluation must take this into account.

Nevertheless, many professional statisticians might profitably read the introductory chapter on "What Statistics Can and Cannot Do." Particularly appropriate is the statement: ". . . Precise looking statistical results may be misleading"; and, again: "The subjective element is very strong and the work of the statistician very often reflects his personality" (page 14).

Even after allowing for the acknowledged oversimplification, one still finds occasional misleading statements. The treatment of the variance is an example. The author explains the relatively infrequent use of absolute deviations to measure dispersion thus: "Many statisticians think we should distinguish in our reasoning between differences above and below the average." Then: "The distinction between above and below differences will cease to be a problem if we put all differences upon a higher level." The variance is defined as the mean of these differences at a higher level.

In this reviewer's experience, the concept of the variance is one which is least understood by laymen. It would seem that the audience to which this book is directed should have a somewhat more precise explanation of why the variance is preferred to the mean of the absolute values of the deviations.

At the beginning of the chapter on sampling, an excellent opportunity is missed to get across the idea that no absolute conclusion about the state of a shipment can be reached from a sample, but that a satisfactory test on a suitable sample from a lot may warrant taking the risk of accepting it.

The book should be of interest to a varied audience. It is essentially popular in nature and easily read (it can be read and understood in two hours), but is not intended as a substitute for a textbook. The description on the jacket seems, in this respect, rather to over-reach itself. It suggests that the book might prove useful as a textbook for a high school course in statistics. As an orientation assignment for the beginner on the first day or so of the course, yes; as a textbook, no!

The Theory of Probability: An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability. Second Edition. *Hans Reichenbach* (Professor of Philosophy, University of California at Los Angeles). Berkeley and Los Angeles: University of California Press, 1949. Pp. xvi, 492. $12.50. English translation by Ernest H. Hutten and Maria Reichenbach.

REVIEW BY JOHN H. SMITH
Department of Statistics, The American University

THIS is a revised translation of the 1934 German edition of a book dealing with philosophical problems related to probability. On the whole, the presentation of these problems is careful and worthy of study.

Reichenbach defines probability (p. 68) as "the *limit of a* [relative] *frequency within an infinite sequence,*" and goes on to say (p. 69) that "no further statement is required concerning properties of probability sequences. In particular, . . . randomness need not be postulated." Although he refers to the theory of R. von Mises which requires independence between any possible rule of selection and the attribute under discussion, Reichenbach specifically excludes independence from his own definition of probability when he says (p. 70) that "probability sequences may be regarded as . . . sequences defined by a rule" giving as the first example "an infinite decimal fraction in which every even number is regarded as the case B."

Reichenbach supplements the "*elementary calculus of probability*" in which "probability sequences are treated with respect to their *external* connections" by a discussion of "the internal order of probability sequences." He considers two probability sequences both of which "go toward the limit $\frac{1}{2}$ of the frequency; however, the sequence (1) shows a strict order while the sequence (2) is irregular, that is, is a random sequence." He goes on to say (p. 131) that "sequences of type (2) will also be called *normal sequences.*" Nevertheless, his formal definition (p. 144) states that "*A sequence B is a normal sequence if it is free from after effect and if the regular divisions belong to its domain of invariance.*" This differs from the collective of R. von Mises for which the probability must be invariant with respect to any method of selection of a subsequence "that can be defined by a rule." In a later section Reichenbach discusses "the fact that a definite probability exists for each individual element." In this connection he uses the device of a "probability lattice," an infinite class of sequences in which each row is a "normal sequence in the wider sense" while the i-th column represents a sequence made up of the i-th elements of the sequences which appear in the rows.

Since this discussion is crucial for anyone comparing alternate definitions of probability sequences, it is in order to consider the clarity with which the author explains the way the columns of the probability lattice are constructed and how the concept of probability is defined as the limit of relative frequency in an infinite sequence. The author refers (p. 168) to playing with an adjustable roulette wheel "with probabilities p_i changing from one throw to

the next so that they converge to the limit p," and goes on to say (p. 169) that "if we keep to a sector width once adjusted and continue playing, the result for the sequence will be a certain limit that differs from the limit of the sequence mentioned first." He continues with the remark that "we could produce all horizontal sequences by playing with the adjustable roulette wheel according to the procedure mentioned; the probabilities p_i would then change from element to element and approach a limit p." In his definition of a lattice he does not state explicitly that the idea of the probability lattice implies that in each horizontal sequence the adjustment of the wheel for the i-th throw must be identical although he suggests this requirement in the statement (p. 164) that the procedure of changing adjustments would "show up in the limits of the vertical sequences since these limits would be equal to p_i." Nevertheless, in later discussion he says that "every adjustment of the roulette wheel determines a sequence, namely, the sequence of throws obtainable by the adjustment. And only because of the interpretation in terms of a sequence can we claim meaning for statements correlating a probability to every adjustment of the wheel." In this latter form Reichenbach's definition of probability seems to be equivalent to stating that the physical conditions determine a hypothetical class or universe of outcomes and that any required probability is the appropriate relative frequency in the universe. This definition would be adequate for the practical applications under consideration. In the opinion of the reviewer, such a definition would have greatly improved the exposition of the earlier part of the book, both by general simplification and by shifting emphasis from properties of sequences to the role of randomness in practical applications. This criticism is to some extent one of choice or order of presentation because the author's presentation leaves little to be desired when he says (pp. 374–5) that "if we are asked to find the probability holding for an individual future event, we must first incorporate the case in a suitable reference class. . . . We then proceed by considering *the narrowest reference class for which reliable statistics can be compiled.* . . . Our knowledge may be improved . . . by the rule: look for a larger number of cases in the narrowest common class at your disposal. . . . There exists only one legitimate concept of probability, which refers to classes."

The treatment of quantitative variables is unnecessarily complicated by the use of unfamiliar terminology. For example, the author calls a variable a "continuous attribute," defines the average as the mean of an infinite sequence, and illustrates (pp. 177–8) an infinite average in terms of the Petersburg problem. He then proceeds to the quadratic dispersion (variance) which he uses in the presentation of Tchebychef's inequality. He also shows that "the quadratic dispersion is additive for independent quantities" and goes on (p. 195) to show (incorrectly) that this principle "can be derived only if the amounts are assumed to be independent." Actually, absence of linear correlation, a weaker condition, is sufficient so that independence is not necessary. Of the many similar misuses of the term *only*, this particular one is perhaps least surprising in view of the statement (p. 236) that "the method

of least squares . . . is too complicated mathematically to be dealt with in this book."

Considerable emphasis is placed on the relationships between probability and induction not only in the last chapter entitled "Induction" but throughout most of the earlier chapters. In this connection, much of the discussion is presented in terms of the concepts of inverse probability exemplified by the "rule of Bayes" to which many references are made. This discussion contains a number of interesting and novel ideas and is, for the most part, carefully presented. Occasionally the author has difficulties in organizing the principles formulated by other writers under the rule of Bayes. For example, in connection with a simple application of the principle of maximum likelihood (pp. 362–3) he states that "like all inferences by indirect evidence, this inference must be construed as an application of the rule of Bayes, and in a complete formulation it will include a reference to antecedent probabilities." In the section in which the foregoing statement appears, it is not clear whether or not it constitutes a mere misuse of the term *must*. Nearly a hundred pages farther on, it appears (p. 454) that Reichenbach does not agree with a position attributed to Fisher to the effect that "selection of statistical hypotheses is governed by rules other than those of probability," and insists that "the only legitimate system for control of hypothetical assumptions is the calculus of probability in combination with the frequency interpretation; it is the only predictive system that is justifiable." The difficulty here is that the author allows the possibility of an attack on Fisher's approach to one type of problem to divert him from his own contributions to another type of problem. Like the method of least squares, the method of maximum likelihood seems to be beyond the scope of this book.

Introduction to the Theory of Probability and Statistics. *Niels Arley* (Assistant Professor of Physics, University of Copenhagen) and *K. Rander Buch* (Assistant Professor of Mathematics, Denmark Institute of Technology). New York: John Wiley and Sons, Inc., 1950. Pp. xi, 236. $4.00.

REVIEW BY JOHN E. FREUND

Associate Professor of Mathematics, Alfred University, Alfred, N. Y.

THE purpose of this book, as stated by its authors, is to give an introduction to the theory of probability and statistics, with particular reference to the problem of application and interpretation.

The first chapter deals exclusively with the meaning of probability, i.e. with the epistemological foundations of the theory of probability. The authors define the probability of an event A, P(A), as "the physical constant for which the relative frequencies are experimental values," at the same time criticizing the frequency theory of probability, as developed by von Mises and others, on the grounds that "it involves a mixture of reality and model

which has been abandoned in all branches of mathematics in favour of an axiomatic treatment." This criticism of the frequency theory of probability is unreasonable, because, staying in the field of pure mathematics, the problem of "meaning" or "interpretation" is completely irrelevant. However, in order to employ a mathematical model for the description of reality, we have to coordinate to the fundamental elements and operations of the model a corresponding set of physical concepts and operations. The probability concept which is defined by Arley and Buch belongs to the model, being (as they call it) the "true value of the relative frequency." This is sheer mysticism, unless we are told something about the relations between this "true relative frequency" and the experimental frequencies, for which it is supposed to stand. This could be done by defining the probability $P(A)$ as the limit which the average of the relative frequencies will approach, if we have a very large number of such experimental relative frequencies. This is, of course, nothing but a somewhat different formulation of the frequency theory of probability, which the authors seem to desire to avoid. The definition of probability, as it is given in this book, defines probability as an idealized concept, standing, in some fashion, for something which involves the experimental frequencies. Unless the authors do specify what this relation is, their definition is, scientifically speaking, meaningless. This criticism does not, of course, affect the axiomatic calculus of probability, which can be developed independent of the problem of interpretation.

One of the brightest features of the book is the large number of interesting problems, e.g. special distributions, drawn mainly from the field of physics. Besides an axiomatic treatment of probability, the customary study of frequency distributions, and a brief introduction to the theory of estimation, it contains a chapter on what is unfortunately (so the authors agree) called the "theory of adjustment." This theory deals with the problem of the estimation of a number of parameters which are functionally related, e.g. the sides and angles of a triangle. This problem is attacked from the viewpoint of maximum likelihood, distinguishing between an "adjustment by elements" and an "adjustment by correlates," depending on whether the restricting conditions are given in a parametric form or whether they are implicit functions of the parameters.

It is difficult to find a suitable place for this book, which is the translation of the third Danish edition, among the literature of the field of probability and statistics. It is a good deal too technical for the nonmathematician, while, on the other hand, it can hardly be called sufficiently comprehensive for a course in the mathematical theory of probability and statistics.

The Pre-Election Polls of 1948. *Frederick Mosteller, Herbert Hyman, Philip J. McCarthy, Eli S. Marks, David B. Truman (and collaborators).* Bulletin 60, Social Science Research Council, New York, 1949. Pp. 396.

REVIEW BY LOUIS H. BEAN

Economist, Office of the Secretary, U. S. Department of Agriculture, Washington, D. C.

THE debacle of the presidential polls in 1948 may have been a blessing in disguise. Now we know a great deal more about the Gallup, Roper, and Crossley methods than before that great mishap, and, research in the various aspects of public opinion polling undoubtedly has been greatly stimulated.

The failure of the 1948 polls posed two basic problems. One was the fear that public reaction might jeopardize both private and governmental confidence in a tool found useful in many fields. The other was the danger of intensifying a conflict among the polling profession between those who argued for the use of probability sampling and those favoring the quota sampling method (commonly used by commercial polling organizations).

In response primarily to the first concern, the Committee on Analysis of Pre-Election Polls and Forecasts of the National Social Science Research Council set up a staff to conduct extensive and quick studies of data on the polls and the 1948 election. By the end of December 1948 the Committee reported its conclusions and recommendations based on the studies of the Committee's staff.

"The Pre-Election Polls of 1948" contains the staff reports to the Committee as well as the Committee's report. In addition, most welcome documentary material in the appendix includes:

1. State samples and survey data on which the final Crossley polls report was based.
2. Bases for final state forecasts of the American Institute of Public Opinion.
3. Data from Roper surveys.
4. The Washington State Poll and Post-Election Resurvey.
5. The Michigan Survey Research Center Survey and Post-Election Resurvey.
6. The Washington Post Plebiscite Poll.

The main part of the volume is devoted to analyses of the staff based on the foregoing material. These cover the following topics: the actual election predictions, the public presentation of polling results, measuring the error, the cross-sections used in predicting the 1948 elections, the interviewing and questionnaire design, processing, estimating and adjustment of survey data, political behavior and voting, evidence pertaining to last-minute swings to Truman and the undecided voter.

Those interested in the various details of polling will want to read these

staff reports as well as the appendix documents. No attempt, therefore, will be made here to review their contents beyond the general observation that the staffs' contributions are objective, fair and competent.

The Committee came to the following conclusions: the pollsters overreached the capabilities of the public opinion poll as a predicting device; the pollsters failed to apply what they had learned from past errors, and failed to ascertain late campaign shifts; there were two major causes of errors: (a) in sampling and interviewing, and (b) in forecasting the future behavior of undecided voters and the shifts toward the end of the campaign; the sources of error were not new but operated more strongly in 1948 than in 1944 and 1940; the manner of presentation of the pre-election polls led to poor public understanding of the nature of errors residing in the polls; and, finally, that "the public should draw no inferences from pre-election polls that would disparage the accuracy or usefulness of properly conducted sampling surveys in fields on which the response does not involve expression of opinion or intention to act. There are more appropriate methods to check the accuracy of such surveys."

Neither the Committee nor the staff members were able to focus on any one particular difficulty that upset the polls. Judging from the several cases where probability sampling was used in comparison with the results obtained by the quota method, it cannot be said the probability sampling could have saved the day. In fact one gets the impression that the superiority of one method over the other for the particular purpose of predicting elections has yet to be established—if it ever can be. After all, elections often turn on a very few votes and can turn on even one vote. It is inconceivable that either quota or probability devices can be made to measure that fine.

Despite the fact that sources of errors are many and complex, I am inclined to think that there were two chief sources, one recognized and dealt with in this report; the other, not. The first is the late shifts that took place after polling ceased. This is discussed in a chapter limited by the meager evidence. It might have been well in this chapter to review the peculiar features of the 1948 election which were particularly conducive to last-minute swings, such as, the general apathy of millions of voters in their response to the two leading candidates; the indecisions in the South due to the Dixiecrat campaign; the indecisions among voters in the northern metropolitan centers created by the appeal of the Progressive Party; and possibly the last-minute shifts to Truman as a result of the general judgment, based on the polls themselves, of an overwhelming Dewey victory.

A most important source of error not at all examined in this report is the farmers' role in the results. The possibility that the polls may have failed to sample the farm vote adequately is referred to in one or two of the staff reports, but the report as a whole misses a most important regional fact about the election and the state-by-state errors. That feature of the 1948 election I have called "The Green Uprising." Only in the Farm Belt (and West Virginia and Massachusetts) did Truman's percentage of votes cast exceed

those for Roosevelt in 1944 by more than two points. Without this upsurge the losses of the electoral votes in the Northeast and South would have been large enough to elect Dewey.

It is true that the polls have suffered from a systematic bias, but the 1948 errors have a distinct regional feature that should have been looked into. Take the Gallup overstatement of Dewey's strength in the nine Western states (Oregon, California, Nevada, Idaho, Utah, Arizona, Wyoming, Colorado and New Mexico). These errors here fall within a range of 0 to 4 percentage points and average 2.8. In the 17 Northeastern states (east of Illinois and Kentucky), with two exceptions, the errors range between −0.9 and 5.8 percentage points and average 2.5 points. The exceptions are Massachusetts (8.7) and West Virginia (8.8 points), where special election conditions helped the Democrats. But for the eleven contiguous Farm states (Wisconsin, Minnesota, Iowa, Missouri, Arkansas, Kansas, Oklahoma, Nebraska, South Dakota, North Dakota and Montana) and Washington, the overstatement of Dewey's strength ranged between 4.6 and 10.9 percentage points, averaging 6.8.

The concentration of the greatest polling errors in the Farm Belt states, the farm vote—so markedly contrary to pattern, and the many reports before and after the election that farmers were unusually quiet at political meetings and immediately after casting their ballots suggest a fruitful source of error in the 1948 polls that yet needs to be probed.

One final observation I am tempted to make in view of its bearing not only on the 1948 election but perhaps its even greater meaning for the 1950 election. This is the matter of turnout. It is examined in Chapter XIII as a possible source of error and the conclusion is reached that "gross turnout had no consistent effect on the outcome and cannot be held responsible for the errors of the polls." To guard against readers underestimating the meaning of turnout, I wish to point out that turnout in 1948 was about ten million below normal (allowing for population growth and trend in voting participation). Had the total vote been larger, the Democratic percentage might have been greater. The polling errors, therefore, also would have been affected.

The examination of total turnout in relation to party percentage should have gone beyond the mere comparison between votes cast in 1948 and 1944. In the first place, 1944 too was a year of subnormal voting, and the actual vote in 1948 should be contrasted with a "normal" 1948 vote. In the second place the failure to find correspondence between a low vote and results in the Farm Belt states may be due to the fact that it is covered up by the "green uprising." Finally, the clearest evidence of the relation of turnout to results is normally to be found in the Northeastern states, in the changes between mid-term and presidential elections.

We are indebted to the authors (and their collaborators, Leonard W. Doob, Duncan MacRae, Jr., Frederick F. Stephan, Samuel A. Stouffer, and S. S. Wilks) for this report as well as for the polling references supplied in the source material. Even more are we indebted to the polling organizations,

particularly to the American Institute of Public Opinion for its decision to carry on with election polls, thus demonstrating their essential uesfulness and reliability by the amazing successes in 1949 and 1950 election polls in both the United States and in foreign countries. Perhaps the same team of experts should be asked to report periodically on these more recent experiences, a they did on the 1948 experience.

Problems in the Study of Economic Growth. *Universities-National Bureau Committee on Economic Research.* National Bureau of Economic Research, Inc., N. Y. 1949. Pp. iv, 250. $2.00.

REVIEW BY M. M. KNIGHT
University of Califorria, Berkeley

THESE five discussions of economic growth follow another series, with two participants in common, Simon Kuznets and Joseph J. Spengler. Kuznets' paper on measurement (here Ch. IV) is reproduced, with prefatory comments, from the earlier proceedings (Supplement VII, 1947, to *The Journal of Economic History*). The foreword and introductory chapter of the present volume are also his. Spengler's earlier paper on the role of the state is heavily drawn upon here (Ch. III), but is greatly extended and considerably altered to fit the heading "Theories of Socio-Economic Growth." In the second paper, J. M. Clark stresses qualitative differences, suggests limits to the quantitative summarization of such disparate elements, and raises the question whether economic growth has not been heavily influenced if not largely determined by non-economic factors. In the final and longest paper, Edgar M. Hoover and Joseph L. Fisher discuss the study of regional economic growth, mainly within the United States.

Kuznets views economic growth as a sustained increase in total output, "sustained" meaning roughly a quarter-century or more. Various conceptions of total output all imply quantitative statements summarized. The advanced industrial nation is indicated as the principal source of data which serve the purpose. Admittedly firm adherence to such units would limit the time discussed to two centuries or less, the space to little more than a few states around the North Atlantic. It is further conceded that data assembled by such a political unit in its own image present an incomplete and somewhat distorted view of cumulative change in the process of economizing. Hoover and Fisher note that intra-national economic regions are politically defined to an extent which makes them rather arbitrary.

The book is chiefly a mass of criticisms, cautions, and suggestions concerning the historical basis of prediction rather than prediction itself. A specialist in the history of economy would extend the list and attach different weights to the factors, roughly along the lines of Professor Clark's paper. The

loose and misleading term "industrial revolution," with its implications of rather narrowly technical causation, has long been out of favor. John U. Nef's work is one example of a tendency to scrutinize the irreversible process of economizing far back of 1750 for the purpose of explaining the phase called "industrialism." German work on "world economy" has sought to expand the discussion beyond political units treated too much in isolation. Walther Hoffmann (whose name is consistently misspelled in Chapter III of this work) has tried to set up models of growing national industrial economy, starting from England and 1700. Colin Clark, mentioned by three of these authors, approached the problem in terms of national income analysis, emphasized statistically measurable effects rather than origins, and hence limited the period seriously considered to less than a century and a half. Clapham had adopted a comparable model, but had applied statistics in a way inviting criticism.

These papers should prove quite valuable for purposes of applying and improving quantitative analysis. Kuznets' statement (p. 6) that "we must consider the quantitative aspects of economic growth as basic" must be addressed to specialists in mathematical treatment. Surely he cannot mean that qualitative analysis is incidental or secondary. It is to be hoped that statisticians who turn to the long-run history of society will recognize how much one must know, in order effectively to know anything, about this vast subject. On the other hand, there is no denying that quantitative summaries by historians are in large part naive. The stress laid by these papers upon the need for collaboration among specialists is well advised. I suspect that this problem will not be solved by specialists already formed, but must be attacked by means of coordinated and radically changed instruction in economics and history, started well back in undergraduate programs.

Problems in the Collection and Comparability of International Statistics. Papers presented at the Round Table on International Statistics. Conference of the Milbank Memorial Fund, November 17–18, 1948. New York: Milbank Memorial Fund, 1949. Pp. 165. 50 cents.

REVIEW BY P. J. LOFTUS

Chief, Economic Statistics Branch, Statistical Office of the United Nations

THE term "collection of statistics" as used by international statistical agencies is something of a misnomer. The phrase really refers to the process of assembling the statistical data of countries in such a way that each set of national data becomes a component of an aggregate that covers a wider universe. The problems discussed in this volume might, therefore, better be described as problems of the supply of statistical data for the world. It is because the national collectors of the primary data do not look upon their

data as components of larger aggregates that problems of incomparability arise. This theme of supply of world statistics and comparability of components is discussed in eleven papers under six headings: demography, health, mortality, food and agriculture, labour, and income distribution. In addition there is an interesting historical paper by Walter F. Willcox on the development of international statistics, which contains an account of organisational developments from the events which led up to the first International Statistical Congress in 1853 up to date. Discussing the future development of international statistics, William R. Leonard concludes that substantial progress can be expected because of the efforts of countries to improve their statistics and conform with international standards.

Dudley Kirk discusses the availability and comparability of world data on population size and characteristics and reveals that there are few demographic indices which approach world coverage. Forrest Linder, recounting the experience of the United Nations in assembling data for the Demographic Yearbook, presents an interesting set of maps illustrating the geographic coverage of the principal kinds of demographic data. Calvert Dedrick, reviewing the plans for the 1950 censuses of the Americas, illustrates how cultural differences among countries affect their census concepts. Max Lacroix describes the primary collection of migration statistics in many countries and shows how international comparability is reduced in the processes of compilation. International efforts to improve world health statistics are described by Knud Stowman. J. T. Marshall discusses the International Lists of Diseases and Causes of Death and describes the circumstances which led to the adoption of a combined list for both morbidity and mortality. The efforts of the FAO to improve the supply of world statistics of food and agriculture and the problems involved are described by Conrad Taeuber. Robert M. Woodbury describes the International Labour Office's work in the standardization of labour statistics. J. B. D. Derksen describes the uses of size distributions of family income and indicates the principal elements of incomparability in these series.

This volume is described in the foreword by Frank G. Boudreau and Clyde V. Kiser as indicating the great progress that has been made in international statistics in the past century. It was probably inevitable that much of the material should be concerned with the international organisational and administrative arrangements which made this progress possible. Nevertheless the volume will be of interest to the increasing number of consumers of statistics whose demand for data extends beyond national frontiers. Here they may find reasons for our present ignorance of the facts of much of human society and hope for greater knowledge in the future.

Business Cycles: Their Nature, Cause, and Control. *James Arthur Estey*, Ph.D. (Professor of Economics, Purdue University). Second Edition. New York: Prentice-Hall, Inc., 1950. Pp. xiii, 527. $4.50.

REVIEW BY ALVIN MAYNE
Wharton School, University of Pennsylvania

IN revising this readable and carefully written textbook, the author has added a chapter describing the events from 1939 to the middle of 1949, a section on the operation of the International Monetary Fund, and about fifteen pages dealing with various taxation proposals which might lead to the stabilization of the volume of consumption. In addition, there are minor changes which have resulted in improved exposition and a few deletions of irrelevant material. The tri-partite organization of the book into: a) description of past cycles, b) summary of cycle theories, and c) problems of stabilization, is unchanged.

The discussion of the various cycle theories, which is excellent, was left virtually untouched, implying that no new developments in the theory of the business cycle have been produced since the thirties. The chapter on the "Theories of J. M. Keynes" certainly was in need of revision to achieve a more convincing demonstration of the cyclical theories implied in the equilibrium analysis of Keynes.

Aside from the "blow-by-blow" account of business activity since 1939 which is measured principally by means of the more important expenditure components of Gross National Product, the author has provided a new edition of the Cleveland Trust Chart of American Business Activity carried back to 1790 instead of 1831 and forward to 1948 and a brief discussion of building cycles based upon Warren and Pearsons' work in that field. Data on the duration and dating of cycles in the United States from Burns and Mitchell's *Measuring Business Cycles* completes the changes to Part I.

The discussion of the International Monetary Fund and the pre-war stabilization operations gives added importance to the international aspects of stabilization. The inclusion of several pages dealing with the impact of a depression in one country on that of a second country under an international gold standard proves helpful in following the later discussion concerning managed currency. In dealing with the problems of a managed currency which is pegged at too high a level, the author might advantageously have devoted several paragraphs to the recent events which caused Great Britain to devalue the pound. In place of the lengthy discussion of Irving Fisher's "compensated dollar" a somewhat shorter presentation is made of the commodity reserve standard proposed by Benjamin Graham. The sections dealing with flexible taxation clearly bring out the problems involved in attempting to administer a system of taxation which is designed to stabilize the volume of consumption.

The revision of the book leaves unaffected the serious discontinuity be-

tween the first third and the last two-thirds of the book. The first part seems dedicated to showing that cycles vary in length and amplitude. Charts are presented for a few measures of business activity, the faults and virtues of common business indexes are indicated, methods of removing trends and seasonal variations are discussed merely to be forgotten as soon as the reader enters the realm of cycle theory. Ten pages of carefully developed charts and text could bring to the student a picture of the great variation in business activity without the danger of boring and confusing him with a variety of piecemeal statistics—mostly out of date—collected by different investigators. I would rather see the author devote the 80–90 pages, so saved, to the integration of the cycle theories with the increasingly important field of business forecasting which seems to be neglected by academic economists.

The author is inclined to limit the activities of the economic statistician to measuring the duration and amplitude of cycles. He is disdainful of any attempts to test statistically the hypotheses invoked by cycle theorists, saying "We are forced to conclude that, in the present state of our knowledge, neither logical consistency nor statistical verification will enable us to choose with confidence between the important, and to some extent contradictory, theories of the cycle now offered" (page 320). The result is that the econometric studies of Tinbergen and his followers are ignored, the magnitude of the multiplier is conjectured from the proportions of income which is saved and consumed without reference to actual data, and generalities concerning cost-price relationships during various stages of the business cycle (e.g., p. 104) are made without a shred of statistical evidence. It is as though a life insurance statistician were given permission to collect data on average longevity and its variation, but "woe be it" if he were ever to relate these statistics to blood pressure measurements to test the hypothesis used by the company to reject applicants and fix premiums. This is the sacred ground of the theorist.

It must be emphasized that, in spite of having "rubbed a statistician the wrong way," Professor Estey will once again earn the gratitude of students and teachers for having provided this excellent summary of the business cycle theories and the problems which are encountered in trying to control the business cycle.

Finally, the reviewer is curious as to the change in heart concerning the part which the entrepreneur is to play. In the first edition he was cast definitely as the hero while in the second edition he is a "hero (or a villain)" (p. 327).

The Balance of International Payments of the United States, 1946–48. *Prepared in the International Economics Division, Office of Business Economics, U. S. Department of Commerce.* Washington: Government Printing Office, 1950. Pp. viii, 275. $0.55. Superintendent of Documents.

REVIEW BY ROLLIN F. BENNETT
New York Life Insurance Company

THIS is the latest in a series of bulletins begun in 1922, the last preceding issue being "International Transactions of the United States During the War, 1940–45." Both in statistical and interpretive material, the latest bulletin is considerably more comprehensive and meticulous than pre-war issues.

The first chapter traces the course of post-war economic readjustments as reflected in U. S. international transactions. The second chapter analyzes changes that have occurred in each major component of U. S. trade. The third deals with investment-income, capital movements, and unilateral transfers. The fourth reviews all transactions, area by area. Concepts, definitions and sources are explained in two appendixes. Data are presented in 67 tables and 21 charts.

The conceptual framework has been influenced by national-income methodology, with a view to making the balance-of-payments statements as useful as possible to compilers of the national income and national product. It is ironical, therefore, that outlying U. S. territories have never been given uniform treatment, being included in the domestic economy for balance-of-payments purposes, but excluded from national income.

Both text and appendixes contain brief discussions of estimation problems. Estimates of services and capital movements are considered less reliable than merchandise data, but there are no detailed evaluations of items in these broad categories. An overall indication of reliability is afforded by the discrepancy between total debits and total credits, which should balance if there were no errors or omissions; in 1948 this discrepancy was $1,242 millions. This discrepancy seems reasonably small compared with the total estimated receipts of $18,373 millions and total estimated payments of $18,689 millions on all accounts other than gold and short-term capital (the latter being reported only on a net basis).

Estimated gains and losses due to revaluations, etc., on long-term holdings serve as a means of reconciling long-term capital movements with changes in year-end investment position. Although the end-of-year position is presented for both long- and short-term investments, no such reconciliation is attempted in the case of short-term capital, and large discrepancies are left completely unexplained.

Because the summary data are broken down and subtotalled in a great variety of ways, nearly every table is tied in with one or more of the others. The reader might have been spared some confusion if the tables had been presented in more logical sequence, with more liberal cross-reference in footnotes.

Effects of Taxation, Inventory Accounting and Policies. *J. Keith Butters, assisted by Powell Niland* (Graduate School of Business Administration, Harvard University). Cambridge: The Riverside Press, 1949. Pp. xvii, 330. $3.75.

REVIEW BY HAROLD BENJAMIN
New School for Social Research, New York, N. Y.

THIS book is the first of a series to be published by the Harvard Graduate School of Business Administration, dealing with the effects of Federal taxation on business. Based upon extensive field interviews as well as on generally available sources of information it discusses the many different methods of inventory valuation currently used for tax purposes and for general corporate accounting. Of special interest is the finding that practical tax administration leaves the tax payer more latitude, i.e., permits lower values, than the regulations promulgated by the Commissioner of Internal Revenue would seem to imply.

Prior to 1939 tax payers were required to price inventories under the assumption that goods on hand at the year end were those most recently acquired. Goods of identical nature will be valued differently in succeeding years under this, the first-in first-out method, and the different values thus attributed to identical items tend to increase reported profits in times of rising prices and vice versa. Since 1939 an alternative method, last-in first-out, is permitted which more nearly matches a year's income with cost incurred during the year. This improvement is had at the expense of an inventory figure in the balance sheet consisting of several layers, each valued at a different year's price level. The fact that the new method has been adopted only by some companies in various industries, mostly the larger ones, and often only for part of their stock, further complicates balance sheet data. One wonders whether an accounting method should not be devised for general accounting, if not for tax purposes, which revalues a year's opening inventory at the price level prevailing at the year end, rather than adjusting the closing inventory to past levels.

The authors present well substantiated estimates as to the extent to which last-in first-out has been adopted in various industries, and as to the effect on reported profits of the various inventory methods. They ably discuss the complications arising from inventory liquidation, the problem of grouping different items, and the inaccuracies resulting from the use of nationwide price indices in the application of last-in first-out to retail stocks. Their conclusions as to the probable effect of a changed income concept on corporate dividend and wage policies, and on inventory and plant expansion appear sound.

To anyone working with inventory statistics, or with national income or wealth estimates, this volume should prove most useful.

Studies in Population. ***Edited by George F. Mair.*** **Princeton, N. J.: Princeton University Press, 1949. Pp. viii, 169. $2.50.**

REVIEW BY JEAN CLAIRE BOWMAN
California State Department of Public Health

In a short review it is impossible to give a full account of each of the twenty-four papers, summaries of papers or critiques that are included in this book, which sets forth the proceedings of the 1949 annual meeting of the Population Association of America. The book is presented in sections corresponding to the five sessions of the meeting, two sessions on current demographic problems and techniques for solution of these problems, one on resources in relation to population and two on human fertility. An attempt has been made to indicate in a general way what was included in the sections. In addition, the content of a few of the papers which were of particular interest to the reviewer has been indicated.

The first general section, *Application of Demographic Data to Current Problems*, starts with "Total Marital Dissolutions in the United States: Relative Importance of Mortality and Divorce," in which Paul H. Jacobson of the Metropolitan Life Insurance Company uses the fragmentary data available on marriage and divorce to reconstruct recent trends in marital dissolutions. He investigates the roles of mortality and divorce (including annulment), and the variation in dissolutions according to duration of marriage. A discussion of this paper by Samuel C. Newman of the National Office of Vital Statistics follows. Length of working life was next discussed, by Seymour L. Wolfbein, Bureau of Labor Statistics, and A. J. Jaffe, Bureau of the Census. The last paper in this section is a "Report on a Series of Illegal Abortions Induced by Physicians," by Christopher Tietze, M. D., National Committee on Maternal Health.

In the section, *Tools for Demographic Research*, the tools discussed are the current population survey, by A. Ross Eckler, Bureau of the Census; the purpose and value of census tracts and divisions in the state of Washington, by Calvin F. Schmid, University of Washington and Clarence E. Batschelet, Bureau of the Census; and the net reproduction rate. The two papers on the net reproduction rate include a summary of T. J. Woofter Jr.'s (Federal Security Agency) "Recent Proposals for Modifying the Net Reproduction Rate," which was presented in the December 1949 issue of this JOURNAL, and a discussion of Woofter's paper by John Hajnal of the Population Division, United Nations, who concludes, "Perhaps the main consequence of looking at reproduction rates in this light is that too much attention should not be paid to them. . . . The long run implications of a given situation need only be investigated occasionally. For other purposes, various measures other than reproduction rates should normally be used."

In the section, *Resources for the World's People*, John D. Black of Harvard University discusses "Population and Scarce Food Resources." He takes issue with books such as those of Vogt and Osborn, describing these ecolo-

gists as unable to accept man in his environment on the same basis as other forms of life and deal with man as a biological unit competing for survival in his environment. He feels that "we need have no fear in this country of declining levels of food consumption because of past abuse of our lands. . . . The increase of more than a third in agricultural output in this country since before the war has not, in the over-all, reduced the productivity level of our lands in any serious measure, and further increases are easily possible at lesser drains on the land than in the past, and at lower costs." He also disagrees with what he calls the "common trough" approach, stating that we must not think only in terms of balancing world totals of prospective food supply against world totals of prospective population. His conclusion is that the total population of the earth will depend, more than upon anything else, upon the rate at which the different countries, and sub-aggregates within countries, reach the stage at which the arts of production begin to advance faster than the population, and standards of living begin to rise.

Following Black's paper, Richard Bradfield of Cornell University elaborates on the availability of resources to produce food and points out that 90 per cent of the world's food is produced in the countries in which it is consumed; there is, however, a need for cooperation among nations to tackle the bottlenecks of ignorance and poverty. He points out that generally the most optimistic views of the world's potentialities for sustained and increased food production are expressed by the agricultural scientists and the most pessimistic views by the non-agriculturalists. The more pessimistic view is represented in this book by Warren S. Thompson of Miami University who has added a paper to those presented at the meeting. He raises some interesting questions in attempting to avoid abstractions and consider real conditions in given countries. He feels that the application of science to agriculture is not likely to remove the pressure of population on subsistence in the near future for the larger part of mankind.

As an example of management of resources, Gordon R. Clapp of the Tennessee Valley Authority contributes a paper on the TVA.

The section on *Value Systems and Human Fertility* includes "The Reproductive Mores of the Asian Peasant," by Irene B. Taeuber, Princeton University, "Demographic Values in the Middle Ages," by Josiah C. Russell, University of New Mexico, and "The Catholic Value System in Relation to Human Fertility," by William J. Gibbons, S. J., National Catholic Rural Life Conference. Kingsley Davis of Columbia University submits a well-taken critique of these papers, pointing out that the emphasis in the first two papers shifts from values as such to the institutions and norms which presumably embody the relevant values. He also points out that the ethical and religious formulations of specialized groups may have only a negligible influence on the demographic behavior of the masses.

Included in the section, *Future Course of Research in Fertility*, is an ably condensed presentation on "Physiological Factors and Their Control," by Howard C. Taylor, Jr., M. D., Columbia University. The paper following,

"An Approach to Research in Overcoming Cultural Barriers to Family Limitation," by Clarence Senior, Columbia University, suggests a shift of emphasis from the broad outlines of value systems and institutions to a consideration of how people actually behave within these systems and institutions and how those who do not conform may be used to influence the remainder.

"Needed Research on Fertility of Negroes," by Ira De A. Reid, Haverford College, indicates certain directions for research in the field of differential fertility as it relates to race and color and suggests the need for a basic reexamination of the color-class categories used in demographic research as they relate to fertility data. The paper on "The Future Course of Research in Fertility: Measures and Methods," by Wilson H. Grabill, Bureau of the Census, discusses reproduction rates, distributional analysis versus rate analysis, and the 1950 census. Two papers, "Some Aspects of Research in Differential Fertility," by Ronald Freedman, University of Michigan, and "The Psychiatric Approach to Research Interviewing," by Moya Woodside, Neurological Institute of New York, conclude the volume.

BRIEFER NOTICES

Mathematical Statistics. ***J. C. Chaturvedi*** **(Lecturer in Mathematics and Statistics, St. John's College, Agra, India). Educational Press, Agra, India, 1949. Pp. 260. Rs 6/-.**

THIS textbook, which apparently was prepared to be used in a second course in statistics in which some mathematical background and interest could be assumed, contains little to commend it. The author approaches the subject from the more classical descriptive point of view and almost completely ignores the more vital design and sampling approach to statistical problems. In other words, the chapter headings follow the familiar pattern: I. Collection of data, II. Classification and Tabulation, III. Graphs and Diagrams, IV. Measures of Central Tendency, etc. The lack of vitality exhibited by such an approach decreases the usefulness of any text attempting to meet the needs of students at the particular level involved, but this text suffers also from the fact that it is poorly written and printed and contains numerous errors and half-truths. The text needs to be carefully edited so as to improve the English since the grammar and choice of words are very poor. The text does contain some of the standard elementary algebraic derivations and proofs which sometimes are hard to find in print, but on the whole, the edition would suffer the least if it were quickly forgotten.

CARL F. KOSSACK

Purdue University

U. S. Department of Commerce, National Bureau of Standards. **Tables of the Binomial Probability Distribution.** Washington, D. C., 1949. Pp. x, 387. Superintendent of Documents, $2.50.

THIS edition is a photographic reproduction of mimeographed tables originally prepared by the Department of the Army. Table 1 gives the individual terms of the binomial expansion $(q+p)^n$ and Table 2 gives the partial sums.

Table 2 was obtained from Karl Pearson's *Tables of the Incomplete Beta-Function* but employs the notation more commonly used for the binomial distribution, $\sum_{s=r}^{n} \binom{s}{n} p^s q^{n-s}$, instead of Pearson's notation, $I_x(p, q)$, where his $x=p$, his $p=r$, and his $q=n-r+1$. Thus, the expression for the partial sum is identical with $I_p(r, n-r+1)$. Pearson carries his x to 1.00; the present tables carry p only to 0.50, but from the relation $I_p(r, n-r+1)=1-I_{1-p}(n-r+1, r)$ the series is easily completed. Table 1, $\binom{n}{r} p^r q^{n-r}$, was obtained by differencing the partial cumulative sums.

These tables are more extensive than any now published: the individual terms run $p=.01(.01).50$, $n=2(1)49$ and $r=0(1)n-1$ and the partial sums run $p=.01(.01)50$, $n=2(1)49$ and $r=1(1)n$. A more extensive but unpublished table (to $n=150$) of cumulative sums, computed at the Ballistic Research Laboratory, Aberdeen, Maryland, is referred to in the Introduction, which cites ten other tables.

The partial sum (page 244) for $n=24$, $r=22$, $p=.45$ is given as .0000001; it should be .0000021.

MARGARET A. LABADIE

University of Chicago

U. S. Bureau of the Census, **Statistical Abstract of the United States: 1949. (Seventieth Edition.)** Washington, D. C., 1949. Pp. xvi, 1122. Superintendent of Documents, $3.00.

THE preface to the 1949 edition is particularly valuable in its explanation of the relation of this edition to previous and future editions and to the supplement, *Historical Statistics of the United States.* A general revision of each section of the annual *Statistical Abstract* is contemplated in order to integrate these volumes with *Historical Statistics.* This revision has been started with six of the thirty-two sections in the current edition. These have been revised in varying degrees, with the section on Transportation probably the most representative of what will be presented in future issues. Data have been added to the sections on Vital Statistics and Foreign Commerce; Area and Population, Immigration, and Public Lands constitute the remaining three sections revised. Less historical detail is presented—only enough summary data to give a proper perspective to the current data—and current statistics receive greater emphasis.

This edition of the *Statistical Abstract* includes thirty-two more tables than the previous edition, making a total of 1,102 tables. The two final tables

contain the data from 1946 to 1948 for the time series which *Historical Statistics* show only to 1945. Cross-references between the *Abstract* and the *Historical Statistics* are provided in the tabular headnotes of all sections including related data.

MARGARET A. LABADIE

University of Chicago

Guide to New Zealand Official Statistics. Second Edition. *E. P. Neale.* Christchurch, Auckland, etc.: Whitcombe and Tombs, Ltd. 1949. Pp. 149.

THIS guide to the use of New Zealand official statistics will prove to be of definite value. The present guide, enlarged and rewritten from the 1938 edition, is a thorough survey of the available statistical sources. Nine chapters deal subject by subject with separate areas of statistical information, and the author's choice of subject headings has been sufficiently judicious to minimize the inevitable overlapping. The following list of chapter headings points to the wide range of material which is surveyed and evaluated: Meteorological Statistics, Population Statistics, Price Statistics, Trade Statistics, Social Statistics, Production and Consumption Statistics, Finance Statistics, Industrial Statistics, and Indicators of Business Conditions.

Those familiar with the author's earlier edition will recall that the *Guide* does far more than outline the available official material. The author takes the reader behind the scenes, as it were, to examine methods of collection, sources of error, and techniques in construction of the official statistics. Those not intimately familiar with the New Zealand statistical services will find this book an extremely valuable supplement to the government's encyclopedic *Official Year Book.*

JOHN A. AUTEN

Massachusetts Institute of Technology

PUBLICATIONS RECEIVED

Adams, Douglas Payne, ed. **An Index of Nomograms.** Wiley & Sons, New York. 1950. $4.00.

Bloomfield, Arthur I. **Capital Imports and the American Balance of Payments, 1934–1939.** The University of Chicago Press, Chicago. 1950. $6.00.

Bray, F. Sewell. **Social Accounts and the Business Enterprise Sector of the National Economy.** Cambridge University Press, London and New York. 1949. $2.50.

Brown, Jr., William Adams. **The United States and the Restoration of World Trade.** The Brookings Institution, Washington, D. C. 1950.

Central Mortgage and Housing Corporation. **Housing in Canada, A Factual Summary.** January 1950. Ottawa, Canada. Vol. 5, No. 1.

Cochran, William G. and Gertrude M. Cox. **Experimental Designs.** Wiley & Sons, New York. 1950. $5.75.

Daugherty, Carroll R. and Marion R. Daugherty. **Principles of Political Economy.** Houghton Mifflin Co., Boston. 1950. $8.00.

Department of Trade and Commerce. **Private and Public Investment in Canada.** Outlook 1950. Ottawa, Canada. 1950.

Emmet, Boris and John E. Jeuck. **Catalogues and Counters, A History of Sears, Roebuck and Company.** The University of Chicago Press, Chicago. 1950. $7.50.

Food and Agriculture Organization of the United Nations. **Yearbook of Food and Agricultural Statistics—1948, II, Trade.** Bilingual English/French with Spanish notes. 1949. Paper. $3.50.

———. **Yearbook of Food and Agricultural Statistics—1948, I, Production, and 1949, I, Production.** Trilingual English/French/Spanish. Paper. $3.50 each.

Gaston, J. Frank. **Effects of Depreciation Policy.** National Industrial Conference Board, Inc., New York. 1950. Paper. 50 cents.

Greece, High Board of Reconstruction. **National Income and Investment of Greece During the Years 1945–49.** Athens. March, 1950. Mimeo.

———, Ministry of Coordination. **The National Income of Greece for the Year 1947.** Athens. 1949. Mimeo.

Haines, George H **Workbook for Busines and Economic Statistics.** Wm. C. Brown Co., Dubuque, Iowa. 1949. $2.25.

Hawley, Amos H. **Human Ecology, A Theory of Community Structure.** The Ronald Press Co., New York. 1950. $5.00.

Hoffman, G. Wright. **Life Insurance at Work.** Life Insurance Agency Management Assoc., Hartford, Conn. 1949.

Institut National de la Statistique et des Etudes Economiques. **Documentation Economique,** Tome VIII, No. 30, 3e trim. 1949. Paper.

International Monetary Fund Staff Paper. Vol. 1, No. 1, February 1950. $3.50 per year (3 issues).

Jacquemyns, G. **Les Radiesthésistes devant l'Opinion Public.** Parc Leopold, Bruxelles. 1950. Institut Universitaire d'Information Sociale et Economique, No. 1.

Johnson, N. L. and H. Tetley. **Statistics: An Intermediate Textbook,** Vol. I. Cambridge University Press, London and New York. 1949. $4.00.

Klein, Ernest L. **How to Stay Rich, The Story of Democratic American Capitalism.** Farrar, Straus & Co., New York. 1950. $2.75.

London Transport Executive. **The London Travel Survey,** 1949. London, 1950.

Maverick, Lewis A. **Economic Statistics.** 1949. $3.00. Published by the author at Carbondale, Illinois.

Morganstern, Oskar. **On the Accuracy of Economic Observations.** Princeton University Press, New Jersey. 1950. Paper. $2.00.

Moore, Geoffrey H. **Statistical Indicators of Cyclical Revivals and Recessions.** Occasional Paper 31. National Bureau of Economic Research, New York. 1950. Paper. $1.50.

Neff, Philip and Annette Weifenbach. **Business Cycles in Selected Industrial Areas.** University of California Press, Berkeley and Los Angeles. 1949. $4.00.

Neff, Philip, Lisette C. Baum and Grace E. Heilman. **Production Cost Trends in Selected Industrial Areas.** University of California Press, Berkeley and Los Angeles. 1948. $4.00.

Proceedings of a Symposium on Large Scale Digital Calculating Machinery. Harvard University Press, Cambridge. 1948. $8.00.

St. Louis and St. Louis County, Social Planning Council of. **Welfare Trends over Ten Years in Greater St. Louis, 1938 to 1948.** St. Louis, Missouri. 1950. Paper. $1.00.

Scoville, Warren C. **Capitalism and French Glassmaking, 1760–1789.** University of California Press, Berkeley and Los Angeles. 1950. Paper. $2.50.

Shevky, Eshref and Marilyn Williams.

The Social Areas of Los Angeles, Analysis and Typology. University of California Press, Berkeley and Los Angeles. 1949. $4.00.

Ulmer, Melville J. **The Economic Theory of Cost of Living Index Numbers.** Columbia University Press, New York. 1949.

Unesco: International Social Science Bulletin. Vol. I, Nos. 1 and 2 and Nos. 3 and 4. Paris. 1949. 90 cents each.

United Nations. **The Effects of Taxation on Foreign Trade and Investment.** Columbia University Press, New York. 1950. Paper. 50 cents.

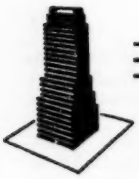
U. S. Bureau of the Census. **Statistical Abstract of the United States: 1949. (Seventieth Edition.)** Washington, D. C., 1949. Superintendent of Documents, $3.00.

U. S. Department of Commerce, National Bureau of Standards. **Tables of the Binomial Probability Distribution.** Washington, D. C., 1949. Superintendent of Documents, $2.50.

U. S. Department of State, Office of Public Affairs. **Patterns of Cooperation, Achievements of International Organizations in the Economic and Social Field,** Washington, D. C., 1950. Superintendent of Documents. Paper.

West Bengal, Government of, Provincial Statistical Bureau. **Report on a Sample Enquiry into the Living Conditions in the Bustees of Calcutta and Howrah 1948–49.** West Bengal Government Press, Alipore. 1949. Paper. Re 1 or 1s 8d.

Whyte, Lewis G. **Principles of Finance and Investment.** Cambridge University Press, London and New York. 1949. $2.50.



McGraw-Hill

NEW BOOKS

INTRODUCTION TO THE THEORY OF STATISTICS

By Alexander Mood, Rand Corporation, Santa Monica, Calif. 431 pages, $5.00

A text for standard courses in statistical theory wih a calculus prerequisite. The author first develops the theory of probability, distribution and sampling, and then proceeds to explore the two major problems of scientific inference: the estimation of quantities, and the testing of hypotheses.

STATISTICAL TECHNIQUES IN MARKET RESEARCH

By Robert Ferber, University of Illinois. 542 pages, $6.00

The purpose of this book is to make available to marketing students and analysts the best and most modern statistical techniques in the simplest manner possible. The author translates mathematics into words to make the book understandable to the student who has never gone beyond college algebra. The most useful and modern techniques are stressed and the author emphasizes those branches of statistics which are of greatest importance to marketing.

HIGH SPEED COMPUTING DEVICES

By Engineering Research Associates. 451 pages, $6.50

Introduces the reader to the general character of computing machines and the arithmetic techniques employed in using them. Shows the basic circuits used to perform various functions in a computing machine, and the relationship between machine capabilities and the required mathematical tools. Intelligible, and up to date, this book will aid the reader in deciding which particular type of computing machine is adaptable to his particular problem.

THEORY OF PROBABILITY

By M. E. Munroe, University of Illinois. In press

Presents in terms of integrals studied in first year calculus an outline of the mathematical framework and a number of the more important results of modern investigations in probability theory. The usual elementary manipulations are covered, but the emphasis is on the concept of a *stochastic variable* and the use of this concept as a framework for the statement and proof of the probability limit theorems.

Send for copies on approval

McGRAW-HILL BOOK COMPANY, Inc.

330 West 42nd Street **New York 18, N.Y.**

Please mention the Journal of the American Statistical Association in writing advertisers